EMISSION ESTIMATION OF HEAVY DUTY DIESEL VEHICLES BY DEVELOPING TEXAS SPECIFIC DRIVE CYCLES WITH MOVES

A Thesis

by

CHAOYI GU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Yunlong Zhang
Committee Members,	Josias Zietsman
	Bruce Wang
Head of Department,	Robin Authenrieth

August 2013

Major Subject: Civil Engineering

ABSTRACT

Driving cycles are acting as the basis of the evaluation of the vehicle performance from air quality point of view, such as fuel consumption or pollutant emission, especially in emission modeling and emission estimation. The original definition of the driving cycle, or drive schedule, given by U.S. Environmental Protection Agency (EPA), is basically a speed-time trajectory which is able to describe the general driving characteristics and driving patterns. Therefore, the development of drive cycles requires a large amount of real data to realize such “generalization”. Then, with such the eligible data collected, it leads to the development of modeling, from traffic modeling to emission modeling, especially for those pollutant emissions which have the public concern.

In this study, focused on heavy duty diesel vehicles (HDDVs), the estimations of the common emissions are being made based on the Texas specific drive cycles, in second-by-second form, collected and generated from five local metropolitan areas, including Houston, Austin, San Antonio, Dallas-Fort Worth and El Paso. First of all, the accurate Global Positioning System (GPS) logging technique is applied for data collection in order to collect not only the moving data but also the relevant geographical information, such as location and roadway, for further analysis. Then, during the progress of data cleaning and data processing, some modifications are made subjectively to improve the deficits of the general methodologies developed by EPA. Afterwards, the specific drive cycles are presented in the format of operating mode distributions, which are also the main part of the input during the emission estimation in Motor Vehicle Emission Simulator (MOVES). Along with all the Texas specific inputs prepared, both the rates and amount of studied emissions are estimated through MOVES. A further comparison is made between the emission rates of default analysis and local analysis to verify the accuracy of MOVES at project level. It is found that the default estimation made by MOVES is accurate for mid-speed cases, at magnitude level. Significant differences happened in low-speed cases and high-speed cases, in which it shows the importance to develop the local drive cycles when estimating the emission rates regionally.

DEDICATION

Dedicated to

My father, Jinhong Gu, and

My mother Xiaoyan Sun

ACKNOWLEDGEMENTS

First of all, I would like to thank my committee chair, Dr. Zhang, who is also my supervisor during the past three years. Without his guidance, I could not finish my thesis and my graduate studies so smoothly. Also, please allow me to express the gratitude to all my committee members, Dr. Wang, Dr. Zietsman and former member Dr. Speed, for the assistance and suggestions from start to the end.

Also, I need to extend my endless appreciation to Reza Farzaneh, who gave me the chance to do the project related to the thesis and Tara Ramani, who gave me the first opportunity to work as graduate assistant with the best colleagues in Texas A&M Transportation Institute. Furthermore, thanks to all my friends and colleagues and the faculties and staffs during my graduate time at Texas A&M University, which is the greatest experience in my life.

Finally, I owe a lot of thanks to my father, Jinhong Gu, and my mother, Xiaoyan Sun, for the greatest love during the past twenty years. In addition, I would like to thank Yana Pan who always stands by my side with the encouragement and support.

NOMENCLATURE

HDDVs	Heavy Duty Diesel Vehicles
EPA	Environmental Protection Agency
HOU	Houston
DFW	Dallas- Fort Worth
SAN	San Antonio
ELP	El Paso
AUS	Austin
MOVES	Motor Vehicle Emission Simulator
GPS	Global Positioning System
CO	Carbon Monoxide
THC	Hydrocarbons
VOC	Volatile Organic Compounds
NO_x	Oxides of Nitrogen
PM	Particulate Matter
TCEQ	Texas Commission on Environmental Quality
TxDOT	Texas Department of Transportation
LOS	Level of Service
FTP	Federal Test Procedure
SFTP	Supplemental Federal Test Procedure
CC	Chase Car

IV	Instrumented Vehicles
EUDC	Extra Urban Driving Cycle
NEDC	New European Driving Cycle
OBD	On-Board Diagnostic
GIS	Geographic Information System
SA	Selective Availability
VTM	Vehicle Miles Travelled
VSP	Vehicle Specific Power
STP	Scaled Tractive Power
HPMS	Highway Performance Monitoring System
ART	Arterial
FWY	Freeway

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1 Background

Driving cycles are acting as the basis of the evaluation of the vehicle performance from air quality point of view, such as fuel consumption or pollutant emission, especially in emission modeling and emission estimation. The original definition of driving cycle (or drive schedule), given by U.S. Environmental Protection Agency (EPA), is basically a speed-time trajectory which is able to describe the general driving characteristics and driving patterns. Therefore, the development of drive cycles requires a large amount of real data to realize the "generalization". With such generalized traffic data, it leads to the development of modeling, from traffic modeling to emission modeling and emission estimation, especially for those pollutant emissions which have the public concern.

Public health is significantly affected by environmental issues during the past century. There are a lot of sources of the air pollutants, one of which is the emission of transportation activities. The vehicle related pollutants are regarded as one important contributor to such negative impact as carbon monoxide (CO), the hydrocarbons (THC), volatile organic compounds (VOC), oxides of nitrogen (NO_x), and particulate matter (PM).

Currently, the most popular emission model in the U.S. is Motor Vehicle Emission Simulators (MOVES), which is developed and released by EPA as the replacement of MOBILE 6. Compared to MOBILE families (MOBILE 1 to 6), preferred

MOVES utilizes a disaggregate approach that enables the users of the model to create and use real drive cycles for the purpose of doing more accurate analysis by the database-driven tool. In addition, it can estimate not only total emissions but also emission rates of various vehicle types (1). In sum, the comprehensive performance makes itself accepted by more and more users and researchers.

1.2 Problem Statement

As a nationwide emission model, a series of national average drive cycles are currently included in the default database of MOVES, which is designed to make sure the comprehensiveness, however, but also creates some deficits in the specific statewide or regional analysis. On the other hand, the emissions and the drive cycle of heavy duty diesel vehicles (HDDVs) that are included in the MOVES model are based on a very limited number of data source whereas they are very important components to the total on-road emission inventory. Thus, focused on the HDDVs, the research effort is designed to utilize the current data source within the state of Texas to provide a new series of drive cycles with a comparison to the default drive cycles. Additionally, the extended output, emissions rates, from the modified database will be compared to estimation rates generated by MOVES default database as well.

1.3 Research Objectives

The goals of this research are firstly to develop the local drive cycles of HDDVs, and then apply them in MOVES to estimate the rates of pollutant emissions during all kinds of movements on various types of roads as the preparation for the comparison between the estimation and real-world on-road emissions measurement in the future. The detailed research objectives are:

- Provide the Texas local drive cycles of HDDVs and the method to develop local ones in the future,
- Provide the foundation to accurately quantify the impact of on-road mobile sources and specifically the heavy-duty diesel sector,
- Provide the basis to develop and quantify emissions reduction strategies of all kinds of movements, such as idle reduction programs in the future,
- Provide TxDOT districts with the ability to more accurately address criteria pollutants,
- Provide districts in Texas with the ability to directly use the knowledge gained about HDDV emission rates and local drive cycles toward investigating fleets such as those found at ports and border crossings.

1.4 Research Benefits

This thesis is based on the project RMC 0-6629. This project will provide the Texas Department of Transportation (TxDOT) with local drive cycles for different regions of Texas with a comprehensive analysis of the emission rates of different vehicle classes and roadway types as well as special cases such as cold start and idling emission rates. Furthermore, the technical and tactical issues of integrating the results of this study into MOVES for formal emissions analyses purposes will be investigated and recommendations will be made based on the findings. The project will also be of great benefit to TxDOT divisions and other agencies such as the EPA and Texas Commission on Environmental Quality (TCEQ).

1.5 Thesis Organization

The thesis is composed of 6 sections. In Section 1, following the background information of emission estimation and emission models, the research problems, objectives, and benefits are briefly introduced. In Section 2, literature related to drive cycles and emission modeling processes are mainly reviewed, such as the general ideas of drive cycles, the existing ones during the previous research, the techniques of drive cycle data collection, the methodologies for drive cycle development and the overview of MOVES. Section 3 presents the content of data collection, including data collection protocols, such as the selection of study areas and vehicles, and the GPS units logging technology. Section 4 illustrates the modified drive cycle development based on the operating mode distribution. Followed by the raw data cleansing, the drive cycles are built step by step as many important amendments added, both on micro-trips and operating modes. Section 5 shows the emission estimation process and the comparison of the estimation rates generated between local drive cycles and default drive cycles. Also, the accuracy and the problems of the cycles are briefly tested and drawn too. At last, section 6 elaborates the summarization of the research work as well as the inadequacies and the direction of future work.

2. LITERATURE REVIEW

2.1 Introduction

In this section, literature related to drive cycles and emission modeling processes are mainly reviewed. In the first part, the general ideas of the drive cycles are given as well as the development during the previous research. Also, the techniques of data collection and the methodologies for drive cycle development are introduced. While in the last part, the overview of the MOVES and its emission estimation process are described in emphasis.

2.2 Drive Cycle

2.2.1 The Drive Cycle Concept

As an important component to quantify the vehicle emissions, drive cycles, a series of data points showing speeds over time, are usually presented as an array in company with external information, such as roadway types and vehicle types. Figure 1 shows an example of such drive cycles developed for an Urban Principal Arterial roadway under Level of Service (LOS) A (2). As it shows, the figure of drive cycles provides the information of a sequence of emission related factors, such as instantaneous acceleration and second-by-second speed. Also, it gives the analyst a way of translating the factors to microscopic emission and from the microscopic level into the macroscopic level by combining the microscopic data captured by the drive cycle.

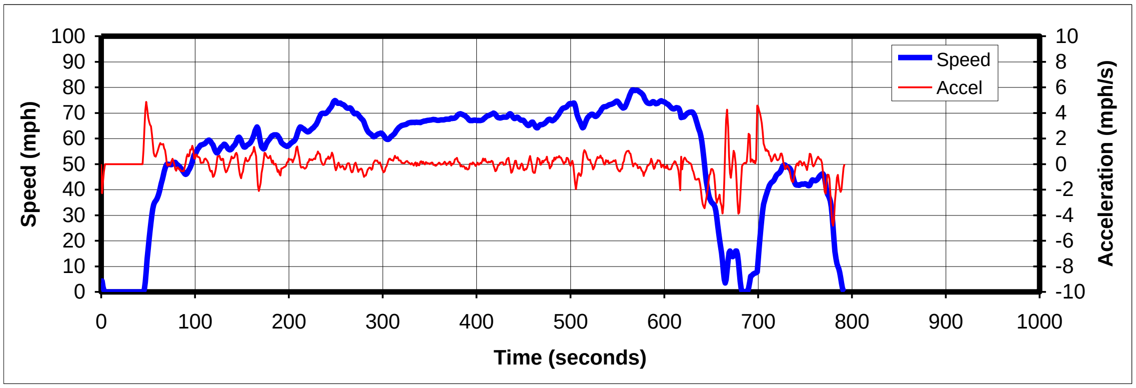


Figure 1. Drive Cycle for and Urban Principal Arterial or Other Freeway/Expressway at Level of Service A

2.2.2 Existing Drive Cycles

Numerous different drive cycles have been researched and developed to describe the grouped driving patterns in various countries, in which they have two general forms, transient drive cycles and modal drive cycles (3). Transient drive cycles are those which involve numerous changes in vehicular behavior over time, such as the frequent speed changes. Modal drive cycles, on the other hand, are generally characterized by protracted periods at a constant speed. The first drive cycles in the U.S., FTP-72 cycles, were developed by Federal Test Procedure (FTP). It was based on data collection in the Los Angeles metropolitan area for use in both vehicle certification testing and emissions inventory development. And then FTP-75 is developed based on FTP 72 but better represents aggressive, high speed driving and the use of air-conditioning. Furthermore, Supplemental Federal Test Procedure (SFTP) drive cycles are further built off of the original FTP cycles by accounting for a wider range of driver behavior including more aggressive driving behavior, high-speed and/or high-acceleration driving and rapid speed fluctuations. In the 1990's, CARB developed the United Cycle for use in California while at the same time the Economic Commission for Europe and the European Economic Community developed the European drive cycles, for example the Extra Urban Driving Cycle (EUDC), the ECE 15 and the New European Driving Cycle (NEDC). Similarly, other international drive cycles are currently used including the 10-15 cycles in Japan, ADC, INRETS, and the Perth Driving Cycle for commercial vehicles (3).

2.3 Techniques for Drive Cycle Data Collection

As illustrated, the basic information contained in drive cycle is the speed and the time. However, when the data collection is taken place, besides the basic information, both the external factors and the instrumental factors which may have the impact on the drive cycles will be captured also. Indeed, there are already numerous technologies have been applied to collect real-time data for the development of drive cycles. Among them, chase cars (CC), and self-contained instrumented vehicles (IV) are the two most widely used approaches.

2.3.1 Chase Car

Both chase cars and instrumented vehicles were applied in MOBILE6 models for speed correction drive cycles. And in MOVES, CC is used in developing the drive cycles for light-duty vehicles (4). Drive behavior is mimicked by following test vehicle in the CC method while laser-base equipment is applied by measuring the distance in between. The contribution of this method could not be denied especially in the early research. However, high cost is the main limitation of the approach when it is applied to relatively small samples. Also, previous research conducted by Moorey, Limanond, and Niermeier showed the potential inaccuracy from equipment, strong variation and the bias between vehicles (5). Even though some recommendations have been pointed out to improve this technology, such as simplifying the route, the data collected by this method still requires a lot of cleaning process since the original data is not directly measuring the speed of the target vehicle (6).

2.3.2 Self-Contained Instrumental Vehicle

As the name indicated, the data is collected from the vehicle itself where the instruments are equipped by this method. The main difference by IV is the direct collection rather than observation so that the bias in observation is eliminated. Also, the additional information is able to be provided by instruments and survey when installing them. Meanwhile, the most outstanding advantage is the data coming from real world as the data is much more influenced by drive only. Additionally, the instruments vary in different forms, such as GPS-based units or vehicular on-board diagnostic (OBD) port connected equipment, which provide more options to match the requirement of the research. Considering the comprehensiveness of the data, GPS-based equipment is heavily preferred in data acquisition as the geographic information is contained as well (6). By this equipment, the location data might be directly tagged in the existing

Geographic Information System (GIS) map, which is relatively a straightforward process by current GIS software. Furthermore, GPS units can automatically record significant amounts of vehicle activity data at the pre-set time intervals they are providing. And the equipment is small, relatively inexpensive and can be easily installed in any vehicles to collect routine activity data without requiring the driver to interact with the unit. However, it needs pay attention that the installation of such units is sometimes strict to the satellite signals when the vehicle is moving or temporarily blocked by the environment. Despite the removal of the deliberate scrambling of GPS accuracy, i.e., selective availability (SA), by the Department of Defense in May 2000 that has greatly improved the positional accuracy of GPS (7), the use of GPS in a moving vehicle will always represent measurement accuracy problems that cannot be completely overcome. Also, the accuracy of GPS receivers has been tested and evaluated by Jackson et al. by comparing the data to the speed obtained from an OBD II ScanTool at second-by-second level (8).

2.4 Methodologies for Drive Cycle Development

When the data collection is finished, the drive cycle development will be the next step. There are three groups of the methodologies that could accomplish this step without any significant drawbacks (9). They are:

- Extrapolation method,
- Selection method, and
- Simulation method.

2.4.1 Extrapolation Method

Constant acceleration phases and reproduced parameters are the characteristics of this method. Both the speed and acceleration are manually cut into different steady phases which are the base of smoothing operating modes got from the method for combining cycles. However, the cycles developed by this method are, kind of easy to follow, not always representative of real driving patterns (3).

2.4.2 Selection Model Method

It is the most widely applied method in developing driving cycles until now. In the approach, the generated drive cycles represent the real time and speed sequences by collected driving characteristics. It tears a big trip to several small parts, which is termed micro-trips, to make the research from macroscopic level to microscopic level. The definition of micro-trip is the key to this method and varies by different researches. Among all the micro trips, the most representative parts will form the drive cycle in which the difference with the entire database has been minimized. Then the drive cycles combined with the survey data are good enough for further analysis, such as emission estimation based on the vehicle types.

2.4.3 Simulation Model

A joint speed and acceleration matrix is utilized in this approach to synthesize the data to a representative drive cycle. Random selection and generation simulates the probability of the sequence of speeds with various characteristics. German drive cycles are developed by this method (3). The method is macroscopic and not that straightforward as the selection model tool and the calculation is somehow a huge work if the database is very large.

2.5 Overview of MOVES Model

2.5.1 Introduction of MOVES

Before MOVES turns out to be widely used, MOBILE 6 is the emission model that was applied by EPA to obtain pollutant emissions and relative factors of various classes of vehicles. In MOBILE model, the predefined database comes from the dynamometer tests and the emission factors are simply multiplied with vehicle miles travelled (VTM) and average speed to get the final emission, which is a kind of aggregate approach (3).

The MOVES, the EPA's latest emission model, follows some characteristics of previous models. First of all, it still utilizes a database centered software framework. Also, it calculates the emission by coupling emission factors and vehicle information together. However, new features are added and successfully functioned. The most important one is that it applies a disaggregate emission algorithm that includes many new external factors for both input and output to provide more options to do analysis at multiple levels. Also, MOVES incorporates energy consumption, geographic information, sort of pollutants, time of the year, vehicle operating characteristics and roadway types together to establish a platform for a more comprehensive analysis than previous. Furthermore, it allows the update of the database so that the flexibility is given to users to catch the change of the characteristics of vehicles' moving and to control local parameters (10).

On the other hand, there are still some drawbacks that cannot be denied. There is a strong agreement in the scientific community that the driving characteristics of each specific area are somehow unique due to different fleet composition, driving behavior, and road network topography. Hence, generalization and specification are hard to co-exist with one default database that is equipped inside of MOVES. Actually, with such a default database equipped, the model just reflects the average at national level even with the option to change the external local factors. Thus, it is less accurate when it is applied to state level or regional analysis. In addition, it is found that the resource of the data which are utilized to form the database is limited in some vehicle activities or vehicle classes, such as the observations for movement of HDDVs (11) (12).

2.5.2 Emission Estimation Process of MOVES

The disaggregate approach in the MOVES is termed Vehicle Specific Power (VSP), which is a measurement of all vehicle factors and non-vehicle factors, such as instantaneous velocity, acceleration and geographical information. VSP is calculated on a second-by-second basis for a vehicle operating over these drive schedules based on Equation 1 for light duty vehicles and Equation 2 for heavy duty vehicles. When it is applied to heavy duty vehicles, it is also termed Scaled Tractive Power (STP) (13) (14):

$$VSP = \frac{A \times u + B \times u^2 + C \times u^3 + M \times u \times a}{M} \tag{1}$$

$$STP = VSP * \frac{M}{f_m} \quad (2)$$

In the equations,

- *u*: the instantaneous speed of the vehicle,
- *a*: the instantaneous acceleration of the vehicle including the impact of the grade (*a = a + sin(a*tan(G/100)*) where *G* is the road grade in percent,
- *A*: a rolling resistance term,
- *B*: a rotating resistance term,
- *C*: a drag term,
- *M*: the vehicle's mass or source mass, and
- f_m: fixed mass factor.

Based on the VSP values and speed values, each observation is categorized into corresponding operating mode bins. The definition of operating mode bins varies slightly among different objectives, such as for emission estimation or for energy consumption. The operating mode bins inside of MOVES for running emissions are listed in Figure 2 (14). Also, this approach adds major flexibility to emission analysis because the emissions of any given observations can be estimated. In MOVES, two different methods are applied to make drive cycles as inputs of emission estimation. One is the operating mode distribution along with the average speed. The other is the original speed profile which is also convertible to the first approach. In fact, two approaches will return almost same results, while the progress is slightly different in synchronizing the drive cycles. Figure 3 graphs the initial emission estimation process of MOVES.

Braking (Bin 0)			
Idle (Bin 1)			
VSP / Instantaneous Speed	**0-25 mph**	**25-50 mph**	**> 50 mph**
< 0 kW /tonne	Bin 11	Bin 21	
0 to 3	Bin 12	Bin 22	
3 to 6	Bin 13	Bin 23	
6 to 9	Bin 14	Bin 24	
9 to 12	Bin 15	Bin 25	
12 and greater	Bin 16		
12 to 18		Bin 27	Bin 37
18 to 24		Bin 28	Bin 38
24 to 30		Bin 29	Bin 39
30 and greater		Bin 30	Bin 40
60 to 12			Bin 35
< 6			Bin 33

Figure 2. Operating Mode Bin Definitions for Running Emissions

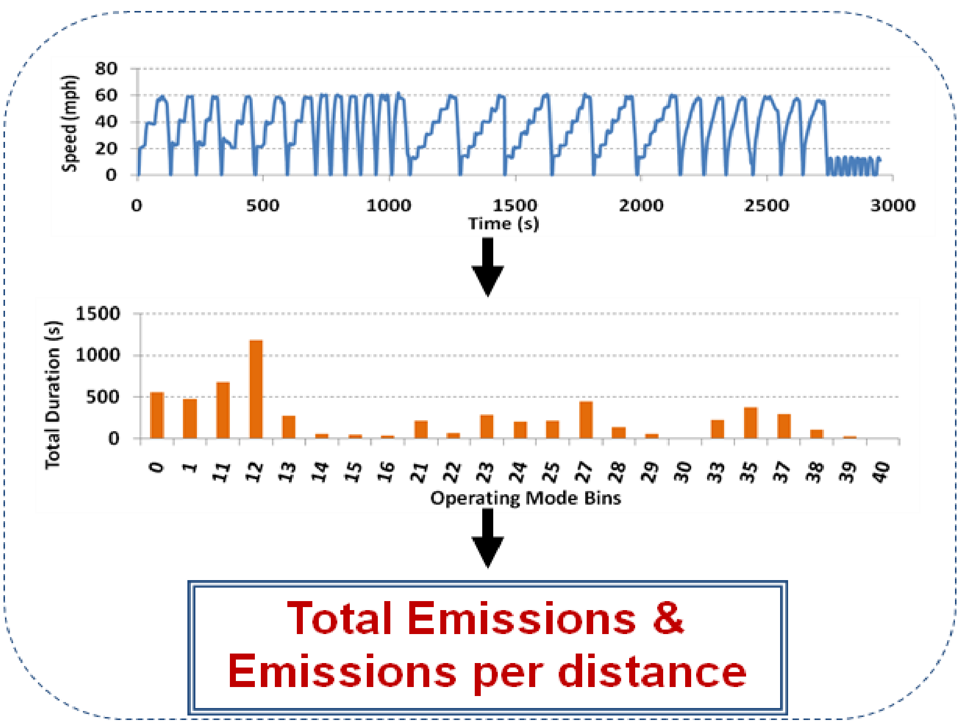


Figure 3. Emissions Estimation Process in MOVES

From a database point of view, there should be really huge amount of drive cycles in order to represent the real world driving activity. To reduce the amount, MOVES applies the average speed as the criteria to group these drive cycles, termed case. Inside each case, there is one representative drive cycle corresponding to the case itself. There are 40 different cases with default drive cycles mapped to specific vehicle types and roadway types that could be accommodated in MOVES. Figure 4 summarizes these cases in MOVES where average speed, speed distribution and the VSP distributions are used to determine the weighting for each drive cycle. As the figure points out, the roadway types and vehicle types are the two main factors in the categorization.

2.5.3 Roadway Types in MOVES

As mentioned above, MOVES simplifies the roadway classification based on the Highway Performance Monitoring System (HPMS) functional classes as shown in Figure 5 (10). Differentiated as rural area and urban area first, the roadways in each group are put into five categories based on their functions in which the definition is broad enough to include all roadway types in general research.

Vehicle Class	Drive Schedule ID	Roadtype	Average Speed (mph)	MOVES Road Type			
				2	3	4	5
Light Duty	1033	Freeway	8.71909	✓		✓	
	1043	Freeway	15.733	✓		✓	
	1021	Freeway	20.6006	✓		✓	
	153	Freeway	30.5	✓		✓	
	1020	Freeway	46.132	✓		✓	
	1019	Freeway	58.7949	✓		✓	
	1018	Freeway	64.3993	✓		✓	
	1017	Freeway	66.3632	✓		✓	
	1041	Non-Freeway	18.5781		✓		✓
	1030	Non-Freeway	25.379		✓		✓
	1029	Non-Freeway	31.0232		✓		✓
	1026	Urban Non-Freeway	43.2662				✓
	1011	Rural Non-Freeway	49.0722		✓		
	1025	Urban Non-Freeway	52.8263				✓
	1024	Non-Freeway	63.66		✓		✓
	101	Freeway & Non-Freeway	2.5	✓	✓	✓	✓
	199	Ramp - Freeway & Non Freeway	34.6	✓	✓	✓	✓
	1009	Freeway & Non-Freeway	73.7991	✓	✓	✓	✓
	158	Freeway & Non-Freeway	76	✓	✓	✓	✓
Medium Duty And Buses	401	Non-Freeway (Bus Only)	15		✓		✓
	402	Non-Freeway (Bus Only)	30		✓		✓
	403	Non-Freeway (Bus Only)	45		✓		✓
	299	Ramp - Freeway & Non-Freeway *	31	✓	✓	✓	✓
	201	Freeway & Non-Freeway *	4.6	✓	✓	✓	✓
	202	Freeway & Non-Freeway *	10.7	✓	✓	✓	✓
	203	Freeway & Non-Freeway *	15.6	✓	✓	✓	✓
	204	Freeway & Non-Freeway *	20.8	✓	✓	✓	✓
	205	Freeway & Non-Freeway *	24.5	✓	✓	✓	✓
	206	Freeway & Non-Freeway *	31.5	✓	✓	✓	✓
	251	Freeway & Non-Freeway *	34.4	✓	✓	✓	✓
	252	Freeway & Non-Freeway *	44.5	✓	✓	✓	✓
	253	Freeway & Non-Freeway	55.4	✓	✓	✓	✓
	254	Freeway & Non-Freeway	60.4	✓	✓	✓	✓
	255	Freeway & Non-Freeway	72.8	✓	✓	✓	✓
Heavy Duty	501	Non-Freeway (Refuse Truck Only)	2.2		✓		✓
	399	Ramp - Freeway & Non-Freeway	25.3	✓	✓	✓	✓
	301	Freeway & Non-Freeway **	5.8	✓	✓	✓	✓
	302	Freeway & Non-Freeway	11.2	✓	✓	✓	✓
	303	Freeway & Non-Freeway	15.6	✓	✓	✓	✓
	304	Freeway & Non-Freeway	19.4	✓	✓	✓	✓
	305	Freeway & Non-Freeway	25.6	✓	✓	✓	✓
	306	Freeway & Non-Freeway	32.5	✓	✓	✓	✓
	351	Freeway & Non-Freeway	34.3	✓	✓	✓	✓
	352	Freeway & Non-Freeway	47.1	✓	✓	✓	✓
	353	Freeway & Non-Freeway	54.2	✓	✓	✓	✓
	354	Freeway & Non-Freeway	59.4	✓	✓	✓	✓
	355	Freeway & Non-Freeway	71.7	✓	✓	✓	✓

*** Not applied to buses**

**** Not applied to refuse trucks**

Figure 4. Default MOVES Drive Cycles

RoadTypeID	Description	HPMS functional Types
1	Off Network	Off Network
2	Rural Restricted Access	Rural Interstate
3	Rural Unrestricted Access	Rural Principal Arterial, Minor Arterial, Major Collector, Minor Collector & Local
4	Urban Restricted Access	Urban Interstate & Urban Freeway/Expressway
5	Urban Unrestricted Access	Urban Principal Arterial, Minor Arterial, Collector & Local

Figure 5. Summary of Road Types in MOVES

2.5.4 Vehicle Classes in MOVES

The primary vehicle classification used in the MOVES model is regarded as "Source Type" for on-road vehicles. The classification roughly corresponds to HPMS vehicle classes as shown in Figure 6. Inside each vehicle type, the detailed description helps the users to assign the operating modes and drive cycles more accurately. As the FHWA's Highway Statistics defined, vehicular classifications in the form of these source types must be used if emissions are to be accurately estimated in the MOVES model. Estimated populations for these classifications must be generated or collected if the estimates for each region are to be accurate (15).

Vehicle Class	Source Type ID	Description
Light Duty	11	MotorCycle
	21	Passenger Car
	31	Passenger Truck: SUV, Pickup Truck, Minivans - Two-Axle/Four-Tire Single Unit
	32	Light Commercial Trucks - Two-Axle/Four-Tire Single Unit
Buses & Medium-Duty	41	Intercity Buses
	42	Transit Buses
	43	School Buses
	52	Single-Unit Short-Haul Trucks
	53	Single-Unit Long-Haul Trucks
	54	Single- Unit Motor Homes
Heavy Duty	51	Refuse Trucks
	61	Combination Short-Haul Trucks
	62	Combination Long-Haul Trucks

Figure 6. MOVES Vehicular Source Types

3. DATA COLLECTION

3.1 Introduction

This chapter mainly introduces the progress of data collection of the study. First of all, the protocols of the data collection have been designed, including deciding the study area and recruiting required vehicles. Also, the method of GPS units has been selected with its accuracy tested. Benefits are show to record the data in this method with not only the required speed profile but also geographic information at the same time. Finally, the data downloaded from the units are stored in certain format for the convenience of drive cycle development

3.2 Data Collection Protocols

Before the data collection is taken place, the protocols for data collection are required to be developed to make sure that the data collected will match the requirements for further analysis in drive cycle development. Since the main purpose of the data collection in this task is to obtain useful speed profile from vehicles, all the possible problems that are related to data collection have to be considered, such as data recording technology, sample size, the duration of data collection as well as the study areas.

3.2.1 Study Area Selection

In order to develop typical Texas specific drive cycles, the research focuses on those large metropolitan areas of Texas. Hence, five large urban areas are selected, including Houston (HOU), Dallas-Fort Worth (DFW), El Paso (ELP), Austin (AUS) and San Antonio (SAN). Also, another main reason behind the selection is that these urban areas generate high volumes of traffic activities and they are also either already in nonattainment or close to nonattainment in air quality for at least one of the critical pollutants. Results in Figure 7 illustrate that the selected metropolitan areas comprise the majority of vehicle activity in terms of VMT as well as the emissions. As presented, each study area is divided in to urban area and rural area by selecting the county with most populations. With the help of GIS software, either the city boundary or county boundary is easy to be determined. For example, the city Austin is belonging to Travis County so that any place located within the county boundary but outside of the city boundary is the rural area of Austin. Table 1 lists the county that selected to represent five metropolitan areas. The drive cycles are expected to be developed for each location type individually.

Table 1. Representative Counties for Study Areas

County\Area	***HOU***	***DFW***	***SAN***	***AUS***	***ELP***
County Name	Harris	Dallas	Bexar	Travis	El Paso

	VMT	VOC	CO	NOx	CO_2	SO2	NH3	PM-10	PM-2.5
		(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)
Texas	723,367,682	1,069,163	13,432,090	2,458,716	841,005,984	17,557	147,116	69,477	43,286
The Selected Metropolitan Areas's Share	63%	57%	59%	50%	58%	56%	64%	55%	53%

Metro Area	VMT	VOC	CO	NOx	CO_2	SO2	NH3	PM-10	PM-2.5
		(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)	(lbs.)
Dallas-Forth Worth-Arlington	175,626,812	228,046	3,180,867	468,647	188,559,201	3,284	36,805	14,460	8,455
Share of the State-Wide Value	24%	21%	24%	19%	22%	19%	25%	21%	20%
Houston	158,457,751	196,485	2,317,919	397,616	171,912,003	3,903	33,114	13,298	7,901
Share of the State-Wide Value	22%	18%	17%	16%	20%	22%	23%	19%	18%
Austin	45,936,097	62,094	746,143	118,024	49,727,989	1,045	9,553	3,916	2,339
Share of the State-Wide Value	6%	6%	6%	5%	6%	6%	6%	6%	5%
San Antonio	56,698,987	93,381	1,127,631	180,376	63,657,385	942	11,639	5,162	3,163
Share of the State-Wide Value	8%	9%	8%	7%	8%	5%	8%	7%	7%
El Paso	16,194,358	31,905	525,536	56,303	17,170,107	693	3,397	1,449	887
Share of the State-Wide Value	2%	3%	4%	2%	2%	4%	2%	2%	2%

Figure 7. Summary of 2008 Emissions and Activity Inventories by On-Road Sources

3.2.2 Vehicle Recruiting

All the vehicles selected in the research are divided into three major classes, light-duty vehicles, medium-duty vehicles and heavy-duty vehicles. However, more than what MOVES has done, the vehicle classification is directly decided by the weight values of the vehicles rather than detailed vehicle classes. The definitions are determined as:

- Light-duty vehicles are those vehicles with weight less than 10,000 lbs.,
- Medium-duty vehicles are those vehicles with weight between 10,000 lbs. and 30,000 lbs., and
- Heavy-duty vehicles are those vehicles with weight greater than 30,000 lbs.

During the recruiting process, for each vehicle, a survey is also done to collect external vehicle information, such as model year, weight, and brand. With this definition, the classification is more straightforward with the survey data collected. On the other hand, even though it is a slightly different classification, any of the 13 vehicle categories in MOVES emission model are still assigned to one of the three categories with survey data and the corresponding drive cycles are applied as well. Also, it is estimated that the sample size is required to be at least five vehicles for each vehicle category in each target area in order to provide enough validated data to cover all the vehicle operating modes for developing Texas-specific drive cycles. With this sample size standard, Table 2 lists numbers of all the vehicles recruited for the entire project in each study area. For the HDDVs, there are enough data for each area obviously.

Table 2. Summary Table of Vehicles Recruited in Project RMC-0-6629

Vehicle Type\Area	***AUS***	***DFW***	***ELP***	***HOU***	***SAN***	***Total***
Light Duty Vehicles	31	11	26	7	8	83
Medium Duty Vehicles	7	15	3	18	7	50
Heavy Duty Vehicles	14	20	34	23	21	112
Total	52	46	63	48	36	245

3.3 Data Collection Technologies

As previous reviewed in Section 2, it reveals that the Global Positioning System (GPS) technology is the best option for collecting speed and location data required for the research at the same time. Also, it was decided in this study that the vehicle recruited are scattered in the different regions of the state. The technology is most applicable since no additional connection to the vehicle is required and the loggers are easy to install.

After examine the accuracy, the proper GPS data loggers that meet the requirements have been selected and applied to do data collection.

3.3.1 Introduction of GPS Units

The majority of current high sensitive GPS units are based on chipsets manufactured by SiRF Technology (SiRF III chipset series) of MTK Incorporation (MTK II chipset series). Following factors are considered in the assessment of units:

- Accuracy of readings,
- Memory size,
- Battery capacity,
- Extension,
- Ease of installation, and
- Accompanied software and other factors.

By comparing all the current commercially available commercial units, QStarz BT-Q1000eX Xtreme Recorder is selected based on the balance of its performance and specifications although it does not meet all the desired specifications. Figure 8 shows the exterior appearance of such units. The Xtreme recorder is based on MTK II chipset with a sensitivity of -165 dB, which exceeds the desired sensitivity threshold. The unit has the capability of recording speed and position data on a second-by-second basis (1Hz) as well as 5Hz and has a memory capacity for 64 hours of observation on the 1-Hz mode. Also, the QStarz unit is equipped with a vibration detector that enables the option of deactivating the unit if no motion is detected for 10 consecutive minutes, which is crucial for save battery power for a longer data collection period rather than wasted when the vehicle is idling. Moreover, it provides an additional battery pack attached to the unit that extends the battery life. During the test for the duration of data collection, it found that around two weeks is the expectation length for each category of vehicles, except for long-haul vehicles which could move more in the same period of time.

Figure 8. QStarz BT-Q1000eX Unit and Extended Battery Pack

3.3.2 Accuracy Confirmation

In order to determine accuracy of the GPS units, a comparison of data has been undertaken. Several trips were made and tracked with both GPS and OBD technology. Both methods included second-by-second speed data, which allowed for calculation of second-by-second acceleration. Three different data sets were investigated. The first set included three distinct trips, the second set had eight trips, and the third set had four trips. As an initial comparison, the second-by-second speed and acceleration profiles of the GPS and the OBD were compared graphically for each data set. It is expected that the speed profiles and acceleration profiles matched up fairly well, indicating that the

GPS and OBD units recorded in fairly similar accuracy in general. In fact, the resulting speed profiles for the first trip are shown below in Figures 9 as a clearer example. As shown, the speeds from the GPS and the OBD follow a very similar trend. On the other hand, there are still some delays or some fluctuations of the data recorded by GPS units by comparing to the data collected by OBD method. The delays mainly happen when there are sudden changes on the speed values. Hence, in this study, the solution is to install three units at the same time in order to minimize such errors as well as the missing data situation if one of them is not working properly.

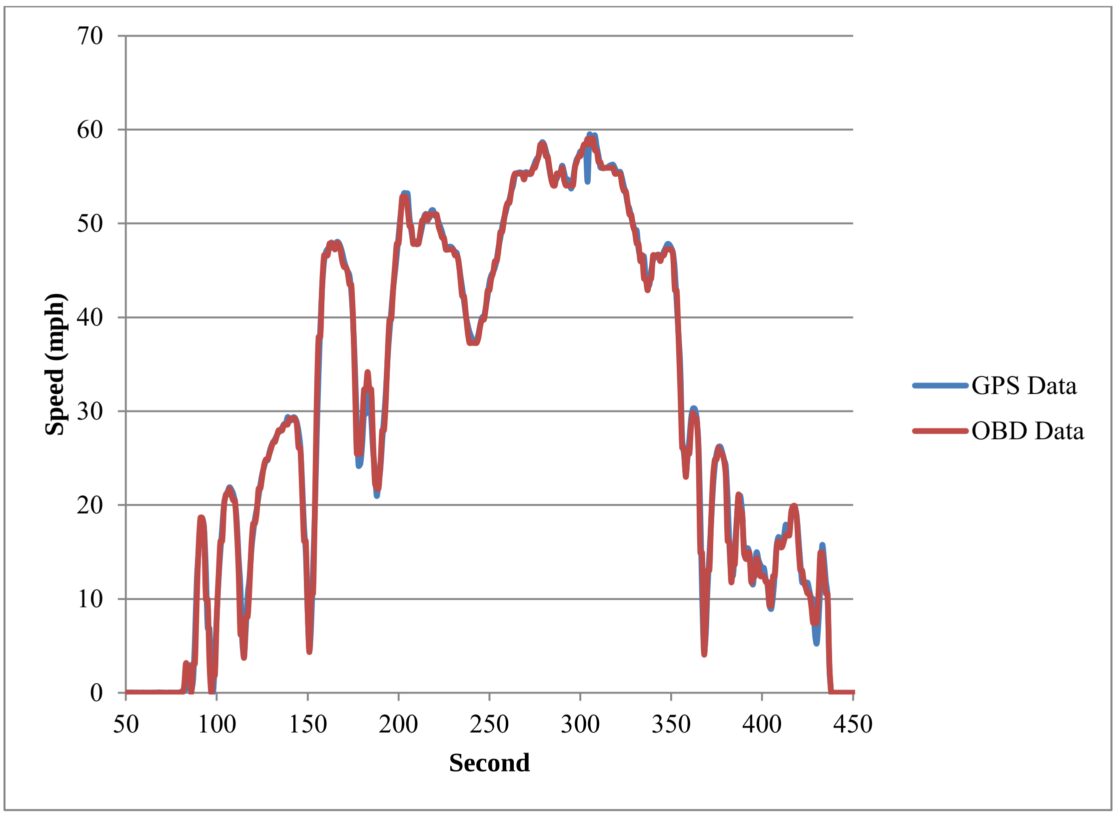


Figure 9. The Comparative Speed Profiles of Sample Trip between GPS and OBD

3.3.3 Data Collection Methodology

In this study, there are two different efforts of the data collection process, planned data collection on predefined routes or on normal routes, where the latter one is the main source of data for the study since the normal activity is desired. The advantage of first effort is able to be under control as desired. It is conducted by TTI stuff member driving a GPS-equipped vehicle on a predefined route in the target area during a selected period of time, including both peak hours and off-peak hours that desired. And a minimum of two days of data collection is taken for this effort for each vehicle. The objective of this effort is to create examples and address the possible gaps of the data for the latter effort. It recruits both individual and fleet vehicles to record normal activity during an extended period of time.

After assembling a pool of potential candidates, the first step would be to have the participant review and a sign a consent form detailing the research being conducted and how their information will be used in the study. Participants are also required to fill out a form giving information about the vehicle make, model, year, fuel type, number of cylinders, and engine size.

Three GPS data loggers will then be sent through the mail or hand-delivered to the participants. Installing three assembles as opposed to one unit in each vehicle is to help to ensure accuracy in the case that one unit is malfunctioning or providing erroneous data. Typically, the devices are set in the driver-side storage compartment (as shown in figure in Appendix B) to ensure that the vibration detector starts recording whenever the vehicle door is opened, which is usually at the beginning of a trip. If the vehicle does not have driver-side storage compartment, the loggers can be placed in vehicle’s glove box or another secure location inside the cabin. The data loggers will be returned after two weeks when the power supply for the unit is expected to be completely drained. For some participants, payment will be issued after the units have been safely returned.

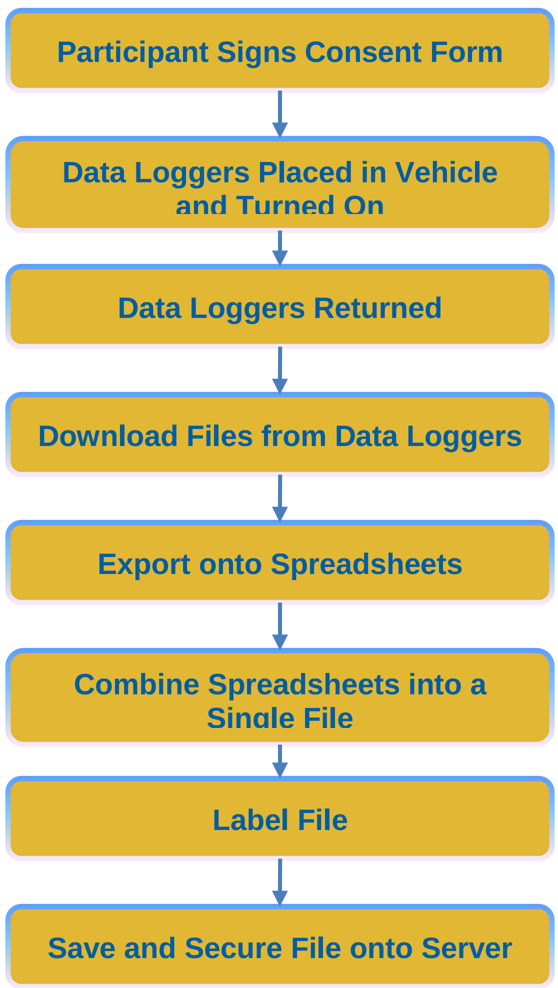


Figure 10. Data Collection Process Flowchart

3.3.4 Raw Data Downloading and Storage

Information from each of the three data loggers are downloaded onto a central server by QSports software and given a unique identifier that will not trace the data back to the original participant. The spreadsheets from each of the three devices are merged into one document that is labeled with variables describing unit number, date of initial activation, and type of vehicle observed. Thus, all the data files are labeled according to following format as “dataLoggerUnitNumber_DownloadDate_VehicleType.csv”, for example, “TTI_10_02102011_SUT.csv”. Figure 10 shows an overall outline of the data collection methodology. For each of the recruited vehicle, both moving information and geographical information are recorded in second-by-second format. And the vehicle ID is added manually as the additional information. Figure 11 shows the parameters which are read directly from the units and from the survey to the drivers. In total, there are 101 heavy duty diesel vehicles recruited with 12,980,975 observations collected.

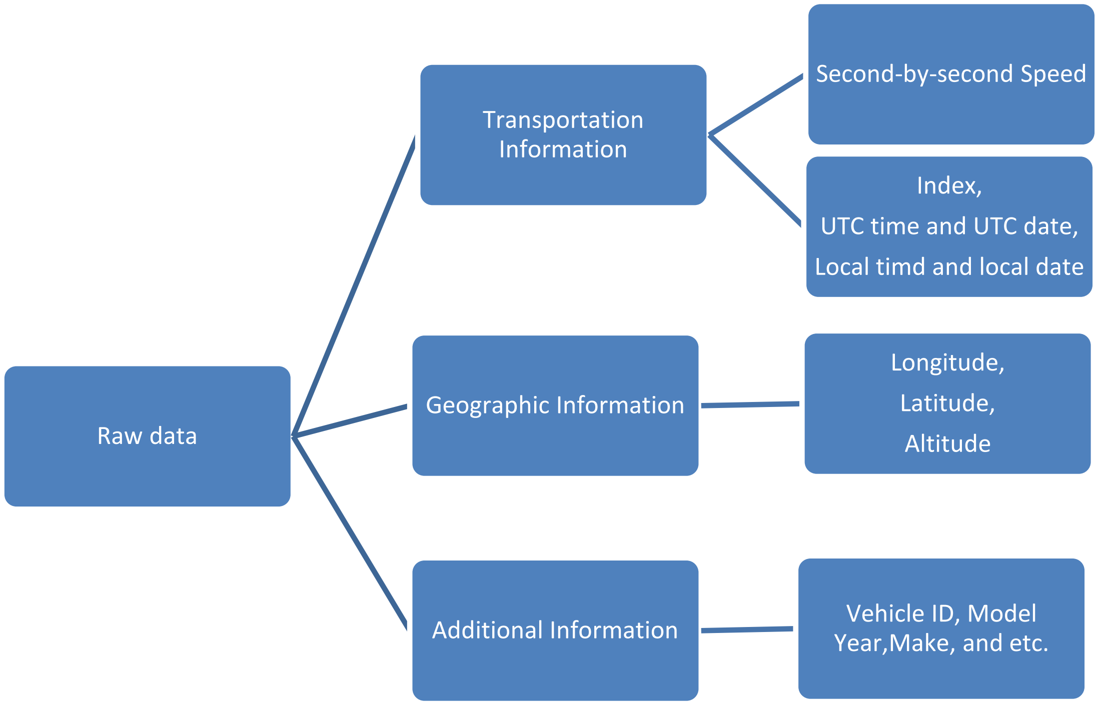


Figure 11. Consist of Raw Data

4. DRIVE CYCLE DEVELOPING PROGRESS

4.1 Introduction

This chapter mainly introduces the modified progress of drive cycle development in selection method. Firstly, an initial cleansing processing of the dataset has been made for the qualification of the generation of the micro-trips. Then, the micro-trips are categorized by area types, roadway types and the bins of average speeds as the cases of the combinations. Furthermore, in each case, the target vector and the representative drive cycle are developed by creating the operating mode distributions.

4.2 Processing of the Raw Data

The ideal situation for the data collected by GPS loggers is continuously second-by-second for the entire battery life. However, as limited by the technology, unusual speeds or missing data happened occasionally, which recalls the importance to do the cleansing of the raw data firstly. Also, only the second-by-second observations cannot be used into drive cycle development directly until they are labeled in an entity, such as trip or micro-trip. As mentioned, developing micro-trips is required to help to format the observation database to micro-trip database. Therefore, as far as applicability and quality was concerned, the following steps were taken place when processing the raw data:

- Do initial data cleansing.
- Convert second-by-second observation database to the form of micro-trip database.

- Flag when the vehicle is in idling situation or approximate idling situation.
- Categorize the micro-trips by roadway types travelled and location with help by geographic information system.
- Bin the micro-trips based on the average speeds.

4.2.1 Initial Data Cleansing of Unreal Observations

After downloading the data from the GPS units, the quality of the data is required to be checked to see if there are any suspect speed values or missing speed values. The check of the suspected speed values is realized by plotting speed values on a plot of acceleration/deceleration versus speed for consecutive observations for each vehicle. Normally, there are mechanical limitations of the acceleration values and deceleration values for the observations which are unreal when beyond them. Suggested by MOVES, the upper limit and lower limit for the acceleration values are as follows:

- The upper limit for the acceleration value is 14 mph/s, and
- The lower limit for the acceleration value is -10mph/s.

However, the thresholds listed above are not enough to rule out the unusual outliers, especially for HDDVs. As far as we know, the upper limit of the acceleration will drop when the speed of the vehicle is higher. For example, an observation with a high speed with a high acceleration value is also unreal even though the acceleration value stays within the limit mentioned above, such as with the speed at 50 mph while the acceleration at 6 mph/s. On the other hand, the deceleration is made by braking which is different with accelerating process. However, high deceleration presents an unusual driving pattern, such as in danger of hitting the vehicle in front which is not the driving cycle that we want include. Hence, mathematically, the negative acceleration, deceleration, will go through the same calculating process of vehicle specific power (VSP) in order to eliminate the observations with such unusual movement. The general upper threshold and the lower threshold of calculated VSP added to the database as:

- The upper limit for the VSP value is 62.5 kW/Mg, and
- The lower limit for the VSP value is -47.5 kW/Mg

In this study, for heavy duty diesel vehicles (HDDVs), the thresholds above are still not quite enough. More thresholds need to be added to the database since the variation of the accelerations is not able to be as large as the light-duty vehicles. However, as emphasized, it is important to divide the decelerating progress and the accelerating progress since the former is quite different with the consideration of loading of engine, which is actually no generation of power when decelerating. Hence, in this research, 99 percentile of all the positive acceleration values from HDDVs database is selected as another upper limitation to decide whether they are qualified. Table 3 lists the 99 percentile value applied in the research.

Table 3. 99 Percentile for Grouped Acceleration Values

Observation Speed Range (mph)	*0-25*	*25-50*	*>50*
99 percentile of Accelerations (mph/s)	3.8	2.5	2

Any observations that appear beyond three limitations above were tagged as invalid observations, although it leads to more separations and break points for the continuous observations originally collected. This problem will be figured out by building up micro-trips.

4.2.2 Micro-Trip Designation

After the initial cleansing, the observation database is now ready for the micro-trip designation. The designations provide the numerous nodes to the observations to be combined as the micro-trips. The micro-trip designation goes after two steps: observations to trips, trips to micro-trips. First of all, if the observations are continuous, they would be designated as one trip. The original time index and trip index are created to label the observations. And next, these trips travelled by all vehicles are separated to micro-trips with new time index and micro-trip index. In this research, the designation is similar to what have done by MOVES in the national research. The start of a micro-trip occurs:

- At the first observation of the entire database,
- When the vehicle has changed,
- When the time step between consecutive time stamps is greater than 1 second; or
- If the current speed is zero while the previous second's speed was not zero.

Each of the micro-trip has its own unique micro-trip identification. In the dataset, the definition of time index is also introduced to the dataset to represent the time sequence of each micro-trip. The beginning of the micro trip is labeled with time index at 0 with an increase of 1 for each second.

4.2.3 Idling Designation

Another kind of suspected observations is the ones with speeds near zero. No doubt the idling emission is one important part of total emissions contributed by vehicles. However, since the purpose of the drive cycle development is focused on the running emissions of the observations, it is really important to correctly label the observations which are idling or approximately idling. It is found that in this data collection method the speed fluctuations sometimes happen even though the vehicles may not move. Therefore, a method is required to be developed to drop those observations from candidate ones. First of all, the study applies MOVES' suggestion to label any observations with the speed less than 1 mph as idling observations. Also, the fact is considered that for a micro-trip, with an average speed less than 1 mph, which means for a one minutes' duration, the vehicle moves approximately at most 26 meters. Also, in this study, the value of 1 mph is below 2.5 percentile, 1.5 mph, among all the average speeds of all the micro-trips. In total, to simplify the process but not losing accuracy, two identifications of such idling situation are developed to deal with the micro-trips:

- A micro-trip that contains 30 consecutive idling observations; or
- The average speed of the micro trip is smaller than 1 mph.

Furthermore, it is found that most of cases of situation 1 are in the micro trips that with the average speed smaller than 1 mph in this research.

4.3 Binning Process of Micro-Trip Database

Each micro-trip has its own average speed. Even when they are coming from one vehicle, the activities are not the same. The binning process is hence created to group the micro-trips with similar activity together based on the average speed values. Meanwhile, external factors, such as research area or road type, which may also have the impact on producing some variances, are added to create more comprehensive classifications. In this research, all the micro-trips are categorized based on the combination of study locations, road types, vehicle types and average speeds of the micro-trips. The term case is introduced to illustrate the combinations. The subsections present all the cases considered in this research as the preparation for the drive cycle development.

4.3.1 Area Designation of Micro-Trips

All the vehicles are recruited at five metropolitan areas in Texas hence the area designations are based on the location where the vehicles are recruited initially by those abbreviations in Table 1. Also, inside each metropolitan area, different roadway resources and the traffic demand reveals in urban or rural area, which are also important to label since the vehicle activities might not be the same.

It is important to pay attention to those boundary-crossing vehicles as well. Also, due to the fleet resource, it is found that there are many area-crossing HD vehicular moving activities, for example from Austin to Houston. Thus, only labeling based on the location where the vehicles are recruited is not a good enough since if treating the micro-trips to assign the area as the resource vehicles, there is a waste of the micro-trips which are out of one area but in another study area. Inside each metropolitan area, it is easy to think about a micro-trip which might cross the boundary between urban and the micro-trip is boundary-crossing in one area, the designation for the micro-trips is based on the location where they start and the majority of them are in. With the help of GIS software, it is easy to make the assignment accurately.

4.3.2 Vehicle Type Designations

As collected, all the vehicles in the database are divided into three vehicle types based on their weights,

- Light Duty Vehicles , with weights less than 10,000 lbs.,
- Medium Duty Vehicles, with weights between 10,001 lbs. to 29,999 lbs., and
- Heavy Duty Vehicles, with weights greater than 30,000 lbs.

Table 2 already lists the distribution of vehicles recruited in this research. And all the HDDVs are labeled as HD in research.

4.3.3 Road Type Designation

Road type is another factor that is taken into account. In MOVES, whether the micro-trip is on freeway or non-freeway is determined by the distance of the micro-trip with the critical value at 3 miles. But, in this research, the approach of geographical information system (GIS) plays an important role in this designation, since the latitude and longitude information for each observation is collected at the same time. With the presence of the observations, the micro-trips are labeled as a freeway (FWY) micro-trips or non-freeway (ART) micro-trips. Here the freeway label means only those access restricted highways. Meanwhile, the distance criterion is still applied. But with the assistance of the geographical information, the critical value is at 2 miles.

4.3.4 Average Speed Bins

Distance travelled, time duration and average speed are the three basic parameters to describe a micro-trip. It is easy to predict that those average speeds deviate randomly for the data that collected. MOVES only provided the cases with speed under 30 mph for urban arterials and the cases with speed over 30 mph for urban freeways of HDDVs. However, when the cases are being developed, the numbers of micro-trips for each bin is the main reason for consideration so that the cycles built from such typical similar vehicle operations will be representative. The binning criterion for cases is set by rounding the average speed to nearest 5 mph for low-speed ART micro-trips and 10 mph for high-speed ART micro-trips and FWY micro-trips for each study area. Table 4 lists the numbers of micro-trips of each study area which proves the sufficiency of the data. And Table 5 lists the final cases which are utilized for drive cycle development in this study.

Table 4. Summary of Numbers of Micro-Trips

Area\Road Type Based Micro Trips		*ART Micro-Trips*	*FWY Micro-Trips*
AUS	Rural	1,671	922
	Urban	3,253	4,476
DFW	Rural	67	353
	Urban	4,121	4,562
ELP	Rural	262	110
	Urban	4,560	1,460
HOU	Rural	746	1,378
	Urban	3,847	2,007
SAN	Rural	293	967
	Urban	4,025	4,165

Table 5. Final Cases of Each Area for Drive Cycle Development

Roadway Types	Cases	Binned Average Speed	Speed (mph) Bin Definition
Non-freeway	A_5	5	[0,7.5)
	A_10	10	[7.5,12.5)
	A_15	15	[12.5,17.5)
	A_20	20	[17.5,22.5)
	A_25	25	[22.5,27.5)
	A_30	30	[27.5,35)
	A_40	40	[35,45)
	A_50	50	[45,55)
	A_60	60	>55
Freeway	F_0	--	[0,5)
	F_10	10	[5,15)
	F_20	20	[15,25)
	F_30	30	[25,35)
	F_40	40	[35,45)
	F_50	50	[45,55)
	F_60	60	[55,65)
	F_70	70	>65

4.4 Drive Cycle Development

As mentioned, drive cycle provides the information of a sequence of second-by-second instantaneous acceleration and speed. It is clear that the optimal solution to describe the driving behaviors in each case is the population of the micro-trips in that case. However, building up a candidate representative cycle is a more convenient solution, even though it is sub-optimal. But, the drive cycle also has its own advantages such as less data involved or user-friendly convenience. Now the question is how to keep the representativeness of such cycles. There are two methods to solve the problem. One is to compare the distribution of speed, acceleration and STP directly second-by second between the entire database and the cycle created. However, it is clear that there are too many possible values of each parameter, still less the three dimensional comparison. The other one is to create the operating modes first to reduce the dimension of the comparison then do so without compromising the quality. The drive cycle will be generated by the second method in this study. The subsections will present the progress of drive cycle development.

4.4.1 General Cycle Building Methodology

There are a lot of methodologies for building the representative cycles by given binned micro-trips. In this study, the operating mode distribution is the way that is applied. The operating mode is the term to describe the vehicle moving behavior based on the speed, acceleration and STP.

First of all, all the observations are binned to the operating mode bins respectively. Now for the entire database of each case, the normalized count of each operating modes will form an operating mode distribution vector, which is called target vector **T.** At the same time for micro-trip number i, it also has a similar distribution vector, which is termed as $\mathbf{M_i}$**.** Similarly, the distribution vector of a cycle is called **C.** The cycle building process is to combine selected micro-trips together while minimizing the value of **T-C.** Here, **T-C** is the value as the evaluation of the representativeness of the cycle where the calculation of is the square root of sum of squared differences. It is easy to predict that a minimized length of **T-C** tends to be 0 as micro-trips are added to build up candidate cycle. The goal of such synchronization is get a **C** enough close to **T** and finish the building when **T-C** is "small enough". Following are the steps of the building process which is showed graphically in Figure 12:

1) Select the micro-trip which has the closet **T-M** value as the first candidate micro-trip vector $\mathbf{M_1}$**,** now $\mathbf{C = M_1}$**.**
2) Select the micro-trip with the closet **T- M** value among the rest of the micro-trips as the second vector $\mathbf{M_2}$**,** now $\mathbf{C = M_1 + M_2}$**.**
3) Repeat step 2 until the **T-C** value is acceptable, and **C** is the sum of $\mathbf{M_i}$**.**

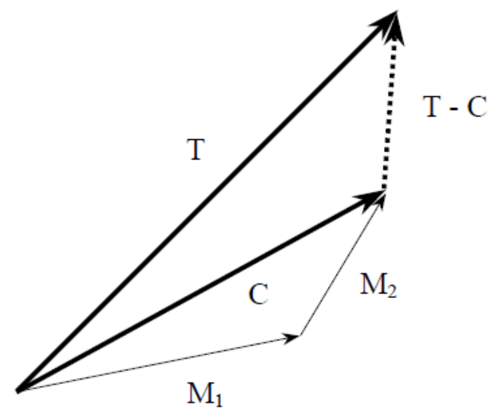


Figure 12. General Cycle Building Process

4.4.2 Cycle Building and Evaluation

Vector in Use

As previously illustrated, the methodology to convert a case of observations to a vector is based on the operating mode distributions. As far as we know, speed, acceleration and VSP are three parameters applied to describe the operations. However, if using three parameters directly to bin the observations, it results in a three-dimension huge vector which is still not that applicable enough. Therefore, the concept of operating mode comes into being to reduce the dimension of the vector to one dimension. Table 6 shows the target vector of case “A_30” in Austin rural area.

Connection between Micro-Trips inside Candidate Cycle

As the definition of the driving cycle, the observations in a cycle will form a continuous movement of the vehicle. Actually, not all the micro-trips are starting and ending at the same speed. There are different gaps between the micro-trips between the end speed of previous one and the start speed of the next one. In spite the general cycle building methodology, the maximum acceptance of the gaps need to be determined or artificial points are required to "fill the gap" to make sure the continuity. In this study, the acceptance of the gap is set at 2 mph, since the artificial points will destroy the realness of the data in some sense. By considering the gap, the micro-trips are then ranked by **T-M** and being linked together to a cycle in the end.

Table 6. Target Vector of Case "A_30" in Austin Rural Area

Operating Mode	Count	Target Vector
0	1444	0.119834
1	505	0.041909
11	401	0.033278
12	611	0.050705
13	243	0.020166
14	196	0.016266
15	158	0.013112
16	166	0.013776
21	1617	0.134191
22	1284	0.106556
23	1617	0.134191
24	1316	0.109212
25	774	0.064232
27	823	0.068299
28	269	0.022324
29	53	0.004398
30	6	0.000498
33	158	0.013112
35	240	0.019917
37	130	0.010788
38	36	0.002988
39	2	0.000166
40	1	8.3E-05

Termination of Cycle Building Process

As mentioned, when the value of **T-C** is "small enough", the cycle building process will be end. How to quantify the "small enogh"? EPA applies the value after 25 steps of cycle creating process as the termination. Figure 13 shows a trend of the **T-C** values as the step of cycle creation goes. However, during the research it is found that such criterion may not provide small enough value of **T-C** sometimes especially when the numbers of micro-trips are relatively low for some cases. Therefore, focused directly on the value, 0.05 is considered as the acceptable value for the **C** as an enough representativeness of the entire database **T**. In total, both two protocols are being applied at the same time to reduce the errors.

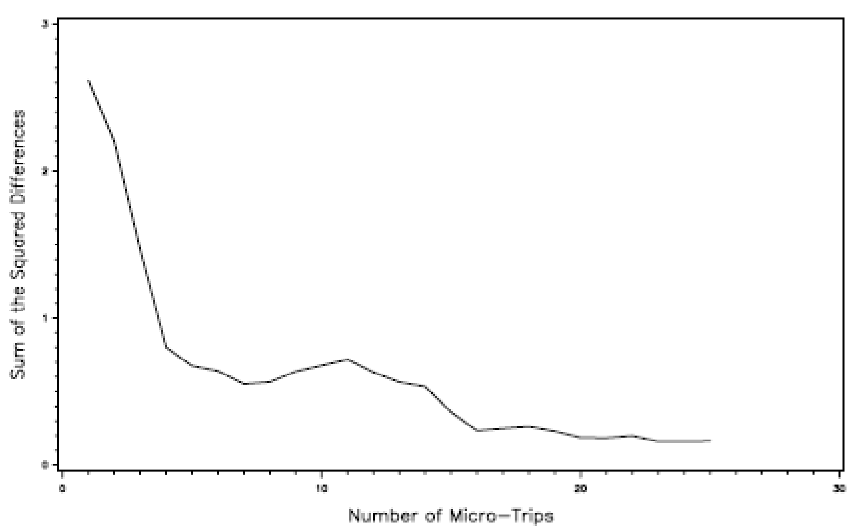


Figure 13. T-C vs. Steps

Modification of Cycles

During the cycle evaluation, only the value of **T-C** seems not enough to describe each operating modes. Figure 14 shows an example of comparison of target vector and cycle vector after the cycle building progress above. Although the value of **T-C** is within 0.05, it reveals the problem of the operating modes above 30 where there are observations in the entire database while empty in the cycle which is questionable to the representativeness. Hence the micro-trips which may not rank in front during the building process while having the missing observations of operating modes above 30 should be added to the cycle, necessarily, still under the previous 0.05 criteria. Figure 15 shows the comparison with the new cycle.

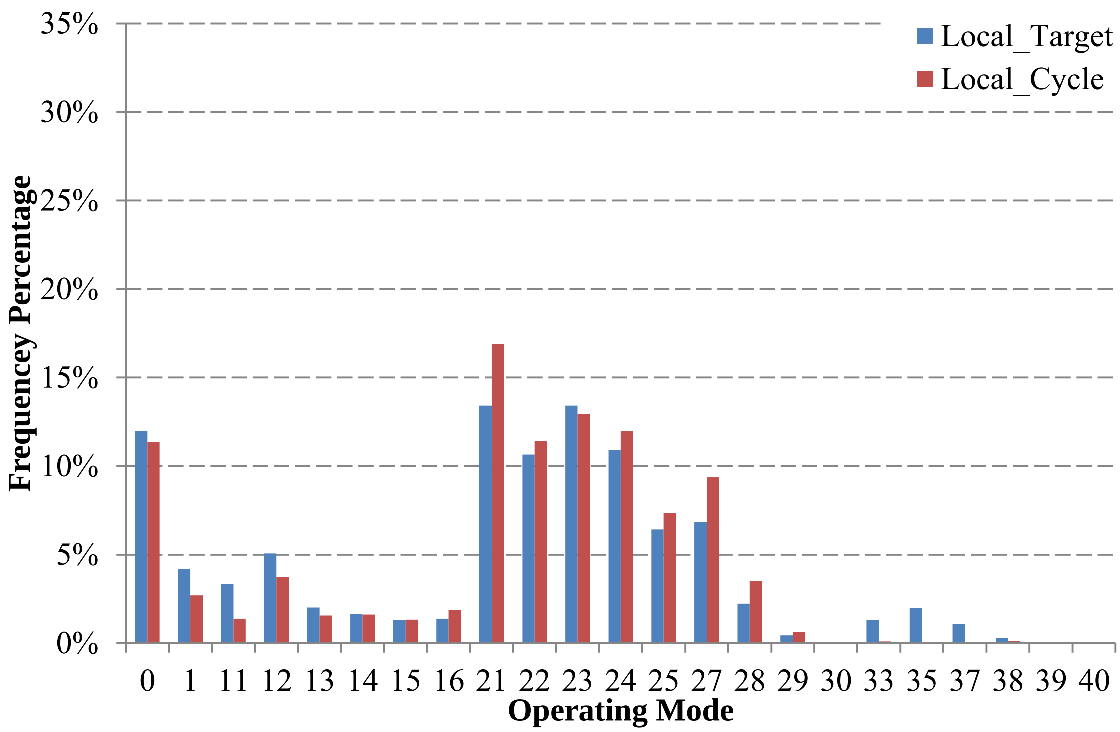


Figure 14. Comparisons of Target Vector and Cycle Vector

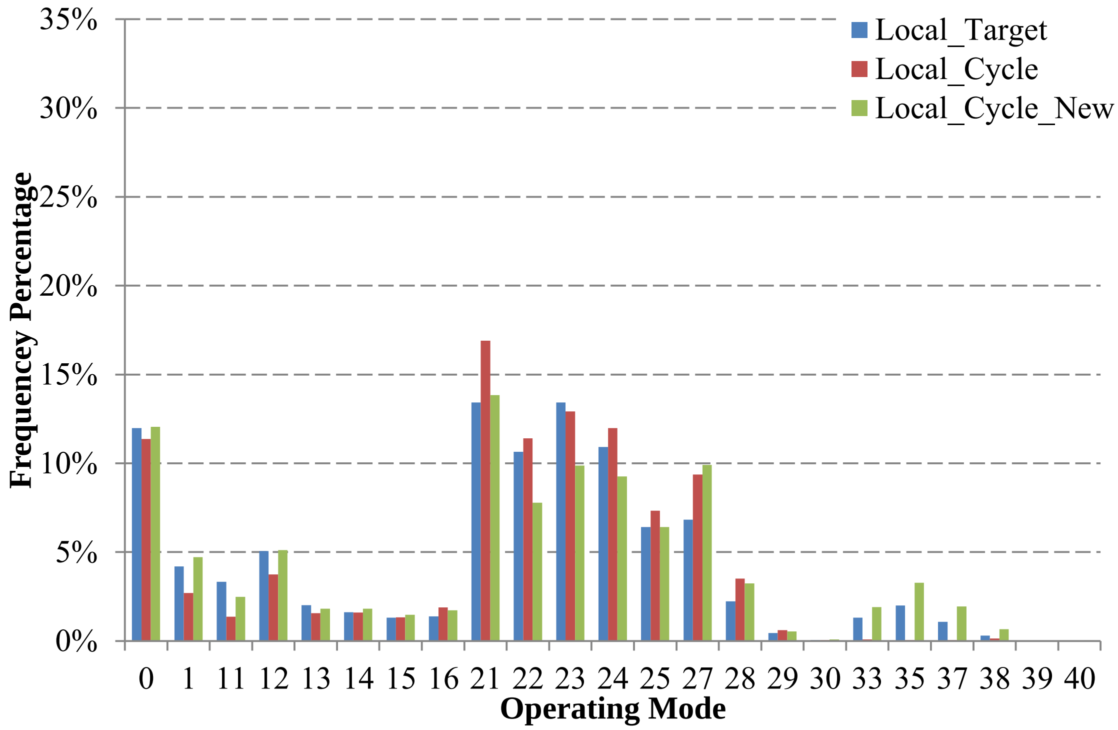


Figure 15. Comparisons of Target Vector and New Cycle Vector

4.5 Results and Conclusions

The previous chapter introduces the data processing progress of drive cycle development. Figure 17 describe the steps of process as a flow chart. Several mainly results are generated after the cycle building process, such as speed and acceleration profiles, operating mode distribution of target vector and cycle vector, and the summary table of each case. These results will contribute in the emission estimation with MOVES and the comparison between MOVES as the inputs since MOVES also has its own default drive cycles. By using “A_30” of Austin rural area as an example, Figure 16 shows the speed acceleration profile between target and cycle. Table 6 lists the basic statistics and Table 7 lists the comparison of operating modes between target and cycle developed. It is clear that the drive cycle developed specific in Texas is different from what MOVES provides as default. The difference will be discussed in next chapter as the part of the analysis of emission estimation and comparison.

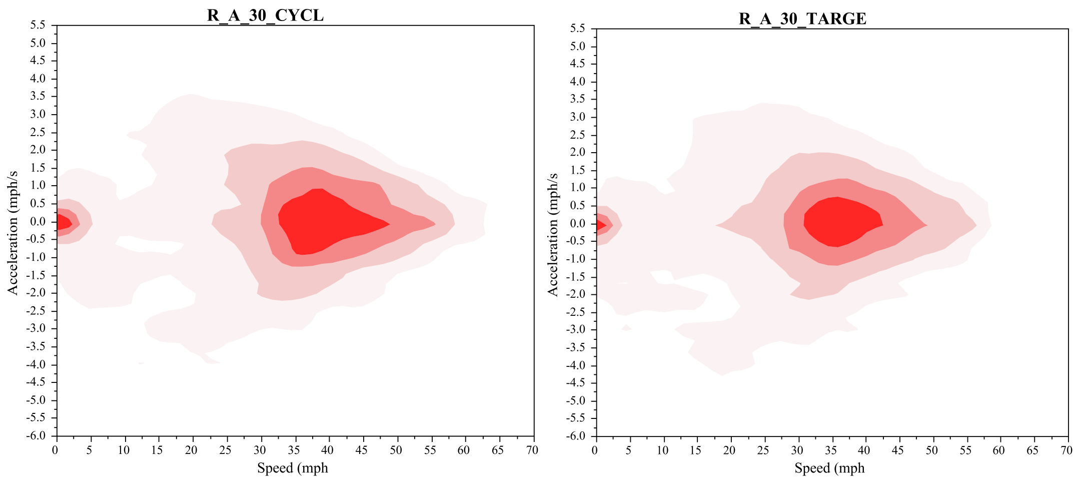


Figure 16. Speed Acceleration Profile

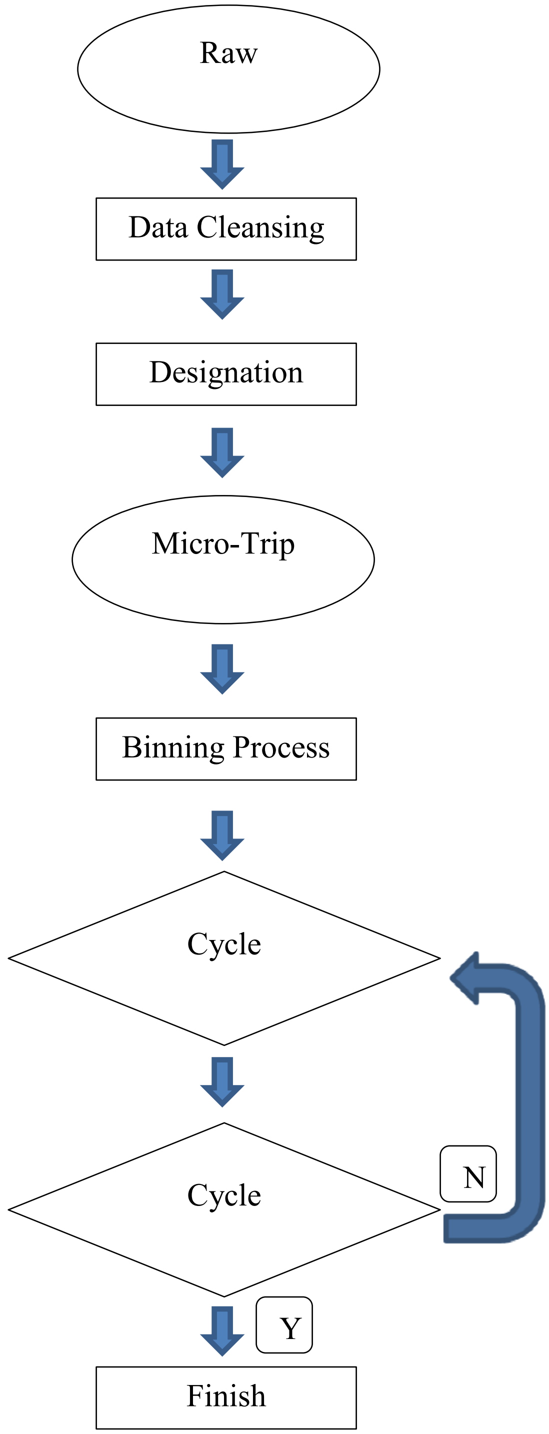


Figure 17. Flow Chart of Drive Cycle Development

Table 7 Comparison of Operating Modes of Case A_30 in Austin Rural Area

Operating Mode Code\Vector	*Target Vector*		*Cycle Vector*	
	Count	Normalized Distribution	Count	Normalized Distribution
0	3,552	0.093518	142	0.108812
1	2,191	0.057685	53	0.040613
11	1,706	0.044916	48	0.036782
12	1,619	0.042625	52	0.039847
13	898	0.023643	44	0.033716
14	812	0.021378	33	0.025287
15	681	0.017929	29	0.022222
16	1,256	0.033068	48	0.036781
21	6,258	0.167462	205	0.157088
22	2,598	0.068840	108	0.082759
23	2,660	0.070033	107	0.081992
24	2,451	0.064531	89	0.068199
25	2,149	0.056379	82	0.062835
27	3,073	0.080907	111	0.085057
28	1,717	0.045206	54	0.041379
29	690	0.018166	21	0.016092
30	389	0.010242	9	0.006897
33	1,332	0.035069	30	0.022989
35	641	0.016876	16	0.012261
37	549	0.014454	17	0.013027
38	378	0.009952	4	0.003065
39	187	0.004923	2	0.001533
40	195	0.005134	1	0.000766
T-M$_{25}$	0.039750			
T-C	0.038776			

5. EMISSION ESTIMATION AND COMPARISON

5.1 Introduction

The section initially presents the emission estimation process of most common gaseous emissions, including gaseous hydrocarbons (THC), carbon monoxide (CO), oxides of nitrogen (NOx), volatile organic compound (VOC), and particular matters (PM). The default analysis made by MOVES generates the default outputs with default drive cycle database. The local analysis goes with the drive cycle database generated in Section 4. Following the results, the differences are discussed from emission rates as well as the drive cycles.

5.2 Emission Estimation Process of MOVES

In the subsections, both the input process and output outcomes are presented completely in order to fully elaborate the emission estimation process by MOVES.

5.2.1 Input Preparation for MOVES

Figure 18 shows the input interface of MOVES. Generally, there are two steps to create an input database, problem information input and template data input. As mentioned, in this research, two different emission rates are prepared to generate. One is the default rate analysis, which is the emission rate estimation by the database equipped by MOVES with only external factors added. The other one is the local analysis where the local drive cycles in the format of operating mode distributions are applied to replace the default drive cycle database. Table 8 shows the inputs that are imported for default analysis and local analysis. Besides the database, the main difference of external factors between the inputs for the two analyses is whether the links are imported with average speeds or operating mode distributions since when only the average speed is provided MOVES will automatically go through its own default database to get the operating mode distributions by aggregating closest default cycles. The following subsections will explain each type of inputs more clearly.

Table 8. Inputs of MOVES for Different Analyses

	Default Analysis	*Local Analysis*
Meteorology Data	Local	Local
Age Distribution	Local	Local
Links	Average Speed Only	Cycle with average speed
Links Source Type	Local	Local
Fuel	Diesel	Diesel
Operating Mode Distribution	Default	Local
Emission Type	Local	Local

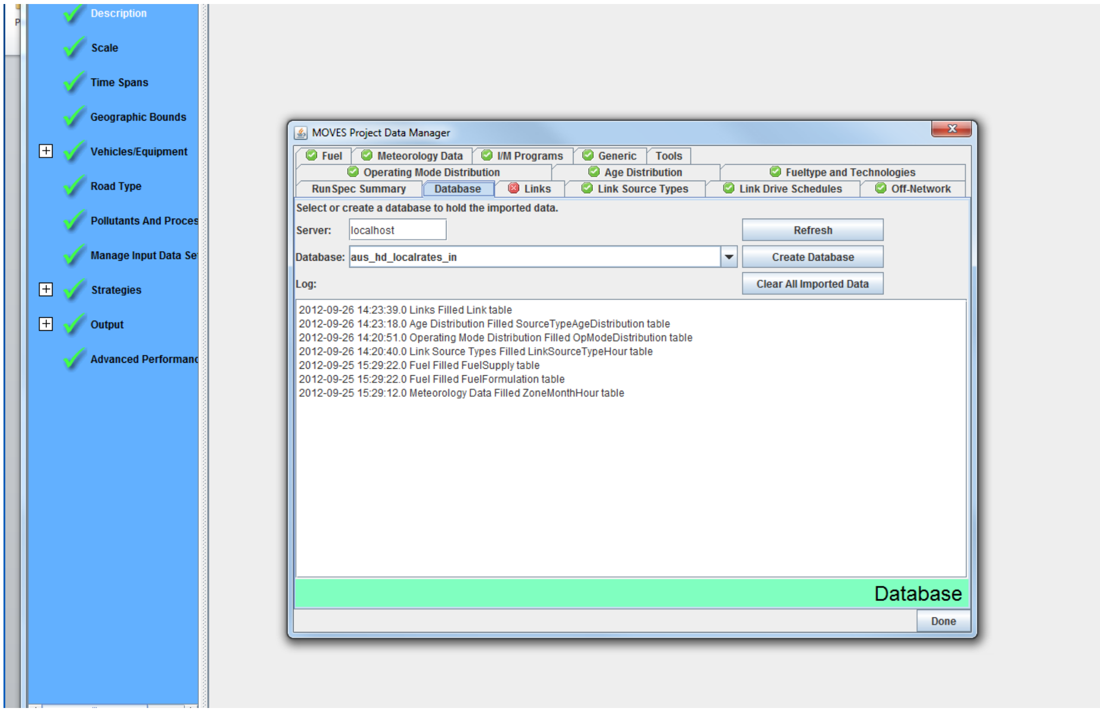


Figure 18. Input Interfaces of MOVES

Meteorology Data

Meteorology data means the external factors which will slightly impact on the emission rate. For each different metropolitan area, and different research time two parameters are considered:

- Temperature, and
- Humidity.

However, since these two parameters vary by seasons, months, or even hours, it is important to set the research time accordingly. In this research, the research time is set at August, 2012, which is fit for time when most of the data have been collected.

Age Distributions

Age is defined as difference between the analysis year and the model year of the test vehicle. For a single vehicle, the age is constant when the analysis year is set. Age distribution is used as an input when the analysis focuses on an area or a roadway segment where there is an age distribution of the vehicles travelled there. For this research, the age distributions are same for two different estimations. Ten years before 2012 are selected as the research model years and the distributions are average for each of the model year.

Links and Link Source Types

Link is the term of drive cycle as the input in MOVES. One link means one completed drive cycle. For the default analysis, only the average speeds of those cycles are imported. MOVES will automatically aggregate the operating mode distributions of nearby drive cycles to the average speeds imported, which are the default inputs as mentioned above. On the other hand, when a link are imported with a completed drive cycle or corresponding operating mode distributions with average speed, MOVES will apply these inputs as the information of links. The meaning of the link source type is the vehicle types moving on each link. For this research, heavy duty is the source type. Therefore, each case will have a unique link associated with the source type of heavy duty.

5.2.2 Output of MOVES

MOVES provides a various types of emissions for further analysis which allow the users to select the typical emissions for research objectives. In this research, critical emissions, such as carbon monoxide (CO), hydrocarbons (HC), volatile organic compounds (VOC), particular matters (PM) and oxides of nitrogen (NO_X), are selected for each study area. The selection is based on the local air quality conditions or non-attainment conditions for a comprehensive analysis. MOVES does not provide the output interface directly. Instead, combined with MySQL, it shows the output information at the format of table. Figure 19 shows the output interface from MySQL. In the output table, following are the parameters provided with the emissions:

- Year ID, which means the analysis year
- Month ID, which means the analysis month
- DayID, which means the day in a week
- Link ID, which is ID of each link when multiple links are imported
- Pollutant ID, which is the ID used for the pollutant analyzed
- Source Type ID, which is the vehicle type
- Model Year ID, which is the year when the vehicle is made
- Road Type ID, which is ID of roadway type
- Temperature, and
- Humidity

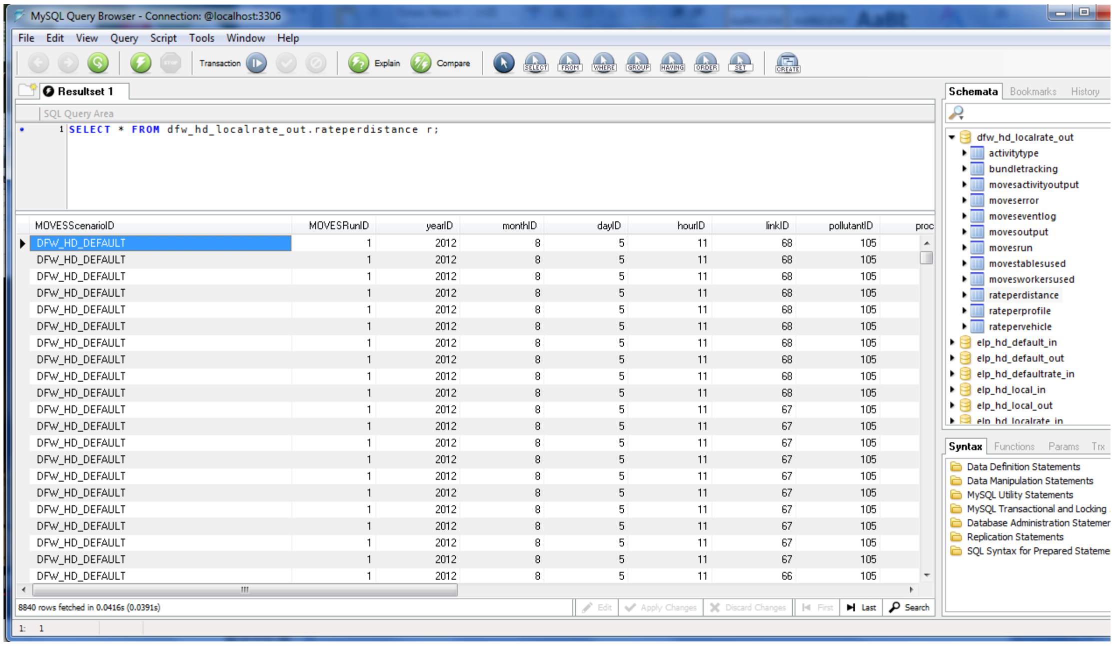


Figure 19. Output Interface

5.3 Comparative Analysis on Emission Rates

The comparative analysis between target vectors is made between local emission rates and the default emission rates for the emissions of CO, HC, VOC and NO_X individually. Brought from the input, following parameters are set and imported for comprehensive analysis for each area:

- Age distribution with 10 years,
- 4 road types,
- 2 area types,
- Temperature and humidity, and
- 17 Cases.

Among all the parameters, in order to simplify the process of analysis, only the model year 2012 is considered in comparison, since the impact of the age is added equally as a factor outside the core calculation.

5.3.1 Preparation of the Comparison

For default analysis, MOVES provide both the emission rate and emission quantity in the unit of emission rate per distance directly, however, only emission quantity is provided for local analysis. It points out the importance to convert them into same unit for comparison. The calculation is shown in Equation 3 as follows:

$$\frac{\frac{Q_{Local}}{P}}{V} = R_{Local} \tag{3}$$

Where

Q_{Local} is the emission quantity with the unit of grams,

P is the population of vehicles with the unit of vehicle per hour,

V is the average speed with unit of miles per hour,

R_{Local} is the emission rate of local analysis with the unit of grams per mile.

To compare the estimation of the emissions, the percentage of difference is set as the criterion. The calculation is listed as Equation 4:

$$Difference\ (D) = \frac{Local\ Rate - Default\ Rate}{Default\ Rate} \times 100\% \tag{4}$$

The control parameter method is applied for the comparison, since it is a four-dimensional analysis with the parameters of emission category, location, roadway types and model years. To simplify the comparison, here the model year is set at 2012 and each of others will be addressed in the selective analyses in following sections. When illustrating the difference and initial analysis based on the emission types, the speed bins IDs are used to replace the actual speeds in this area based analysis to make the graphs more intuitively.

5.3.2 Characteristics of Selected Emissions

As a simulator, MOVES does a lot of calculation to translate the input information to output emissions. Although multiple different types of emissions are selected as the research objectives, there are some relationships between by looking at the calculation progress directly. Also, a substitute way is looking at the output graphs to find such directly since graphs are better understood. In this section, Austin area is selected as the example.

Figure 20 to Figure 25 show the graphs of the differences of emission rates between local and default versus the case studied in Austin area. Here are the findings:

- The shape and location of the plot of CO_2 emission is similar to the one of NO_X emission and the one of PM_{10} emission.
- The shape of the plot of CO emission is unique.
- The shape and location of the plot of THC emission is similar to the one of VOC emission.
- The scale of the plot of CO_2 emission is larger than the one of NO_X emission.
- The scale of the plot of NO_X emission is larger than the one of PM_{10}

emission.

- The scale of the plot of THC emission is same as the one of VOC emission.

Consequently, based on the plots, the selected emissions could be grouped and analyze anyone in the group could reflect the change of all the other members in this group. Obviously, the groups are:

- CO_2, NO_X and PM_{10}
- THC and VOC
- CO

In following sections, THC emission, NOx emission and CO emission are selected to represent the three groups. Also, considering the fact that there are a lot more VMT and emissions generated in urban area than rural area which is also revealed by the numbers of observations collected. Urban area is selected to do the following comparative analysis.

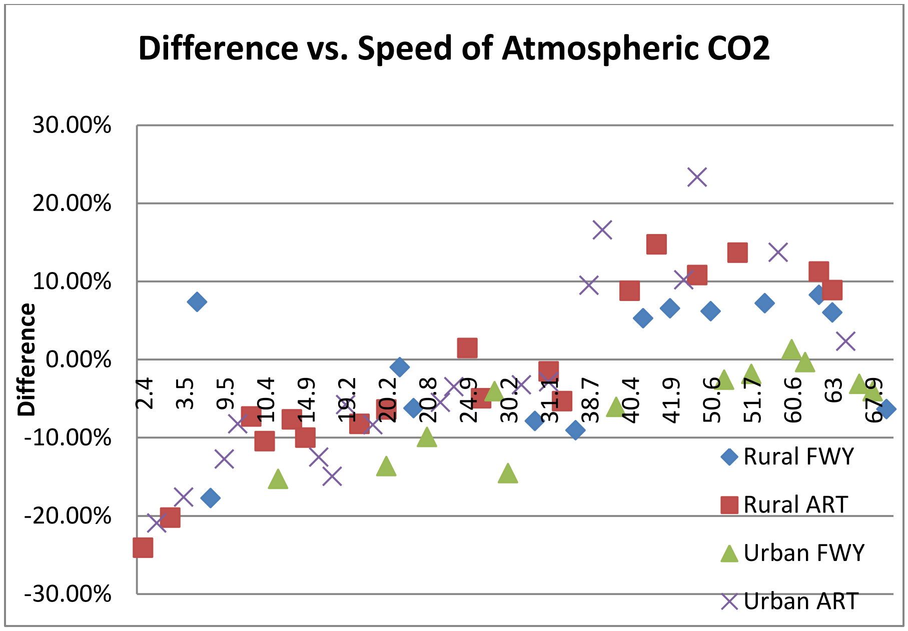


Figure 20. Percentage Difference of CO_2 vs. Average Speed Cases in Austin

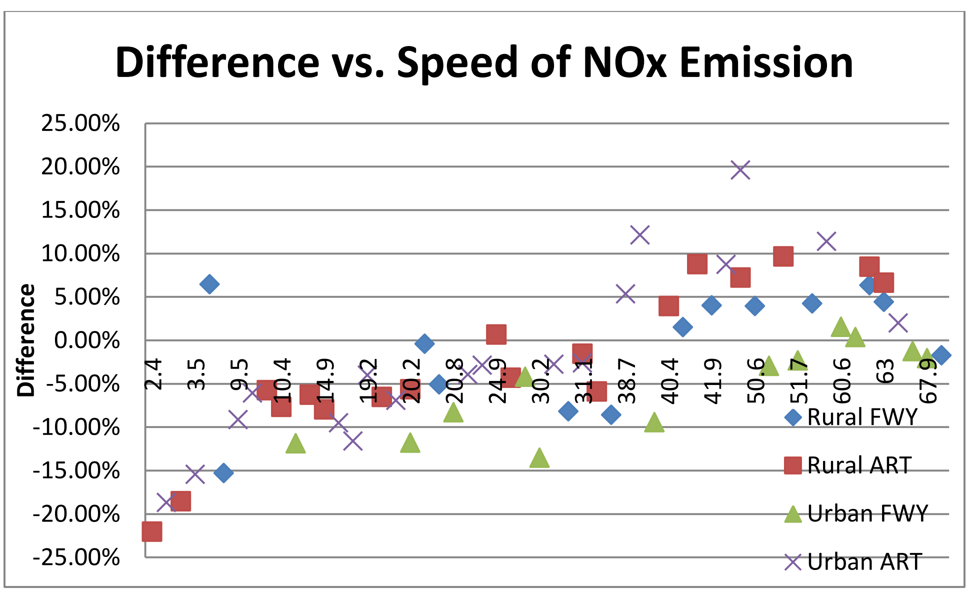


Figure 21. Percentage Difference of NO_X vs. Average Speed Cases in Austin

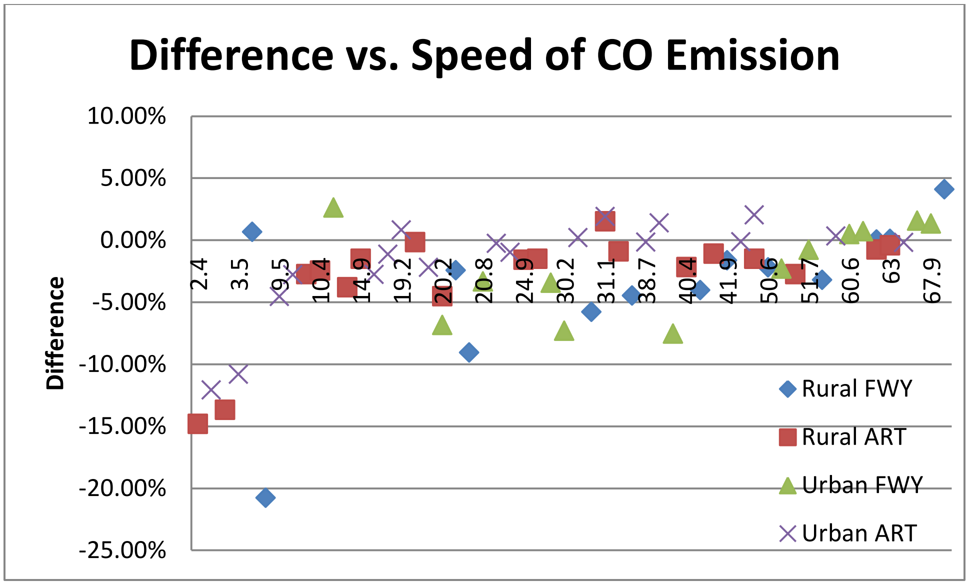


Figure 22. Percentage Difference of CO vs. Average Speed Cases in Austin

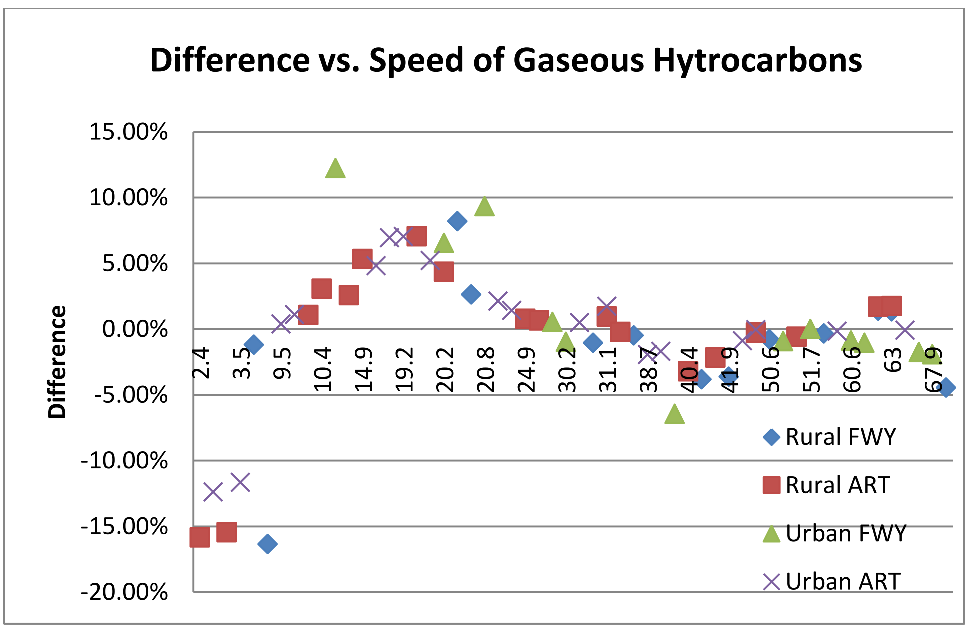


Figure 23. Percentage Difference of THC vs. Average Speed Cases in Austin

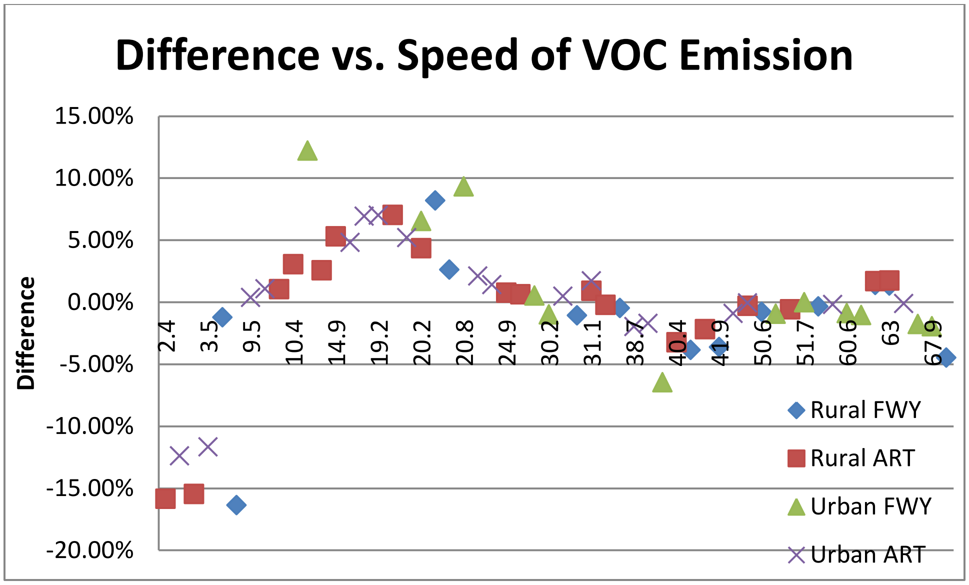


Figure 24. Percentage Difference of VOC vs. Average Speed Cases in Austin

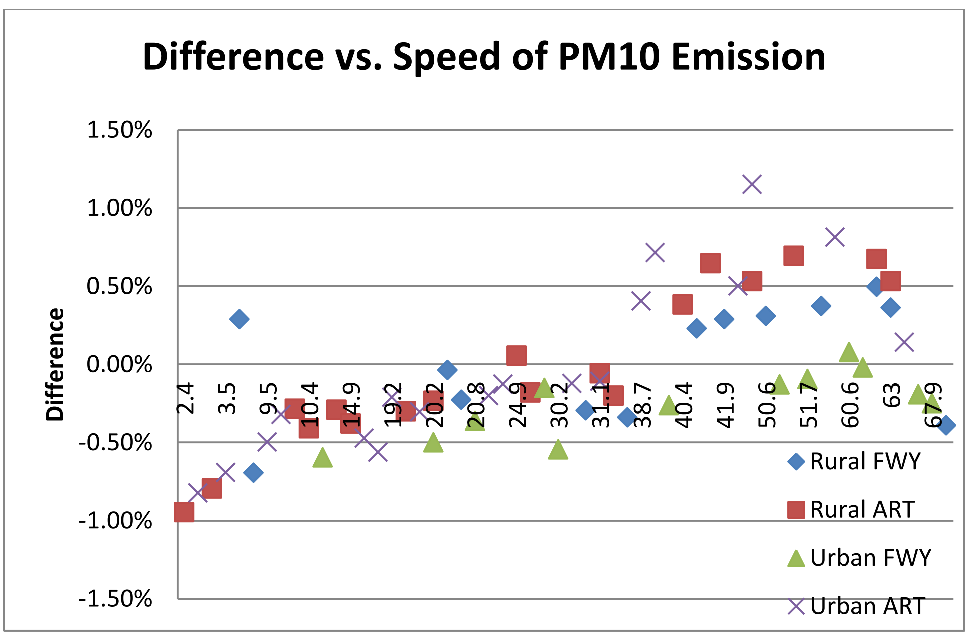


Figure 25. Percentage Difference of PM_{10} vs. Average Speed Cases in Austin

5.3.3 Running THC Emission

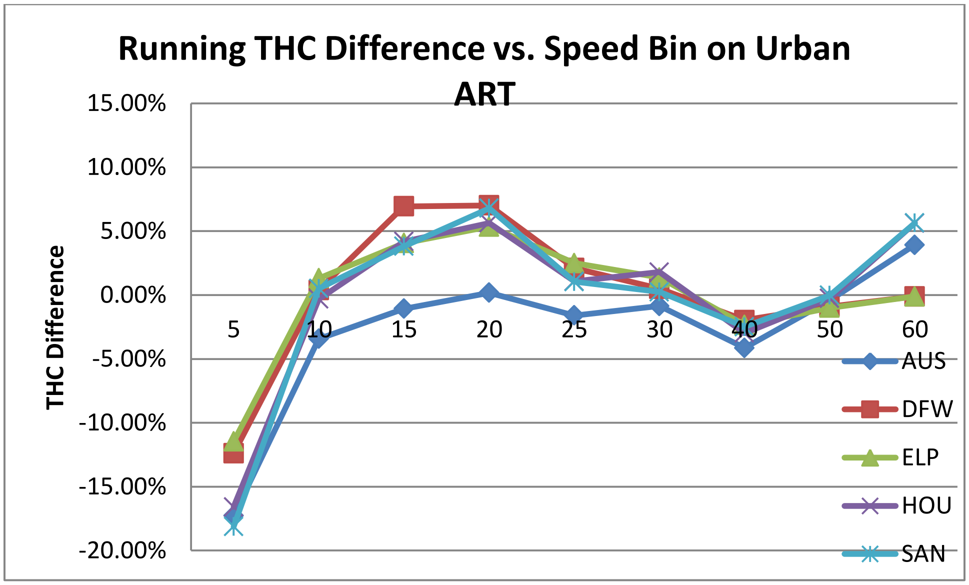
The total gaseous hydrocarbons (THC), completely contrary to the gas emissions after combustion, indicate the total incomplete combustion so that the difference plot should be opposite to the CO_2 emission theoretically. Figure 26 shows the percentage differences of THC emissions as the speed increases of all the five metropolitan areas on either urban arterials or urban freeways. There are several interesting facts that for urban arterials:

- The default results of MOVES overestimate 12% to 17% of running emissions in the case of the average speed is 5 mph on urban arterials.
- The default results and the local results are almost the same in the case of 10 mph, 30mph and 50 mph on urban arterials.
- Between the case of 10mph and 30 mph, the default results of MOVES underestimate up to 5% and in the case of 20 mph, the shape reaches its peak except for AUS.
- In the case of 60 mph, the MOVES underestimate 5% in SAN and AUS.
- In the case of 40 mph, the MOVES overestimate up to 5% also.

Also for urban freeways, the graph shows:

- Similar to urban arterial, the default results and the local results are almost the same in the case of 30 mph and 50 mph. And the underestimate peak happens in the case of 20% in which the value is up to 10%.
- In the case of 40 mph, the MOVES overestimate up to 5% also.
- Unlike the urban arterial, in the case of 10 mph and 70 mph, the differences

vary area by area, and distribute around the horizon.



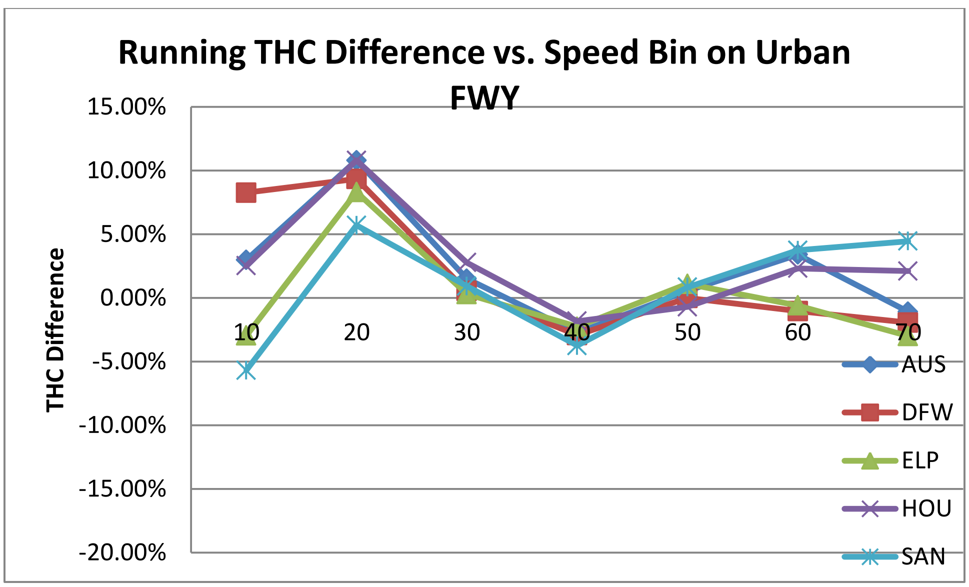


Figure 26. Distributions of THC Differences of Five Study Areas

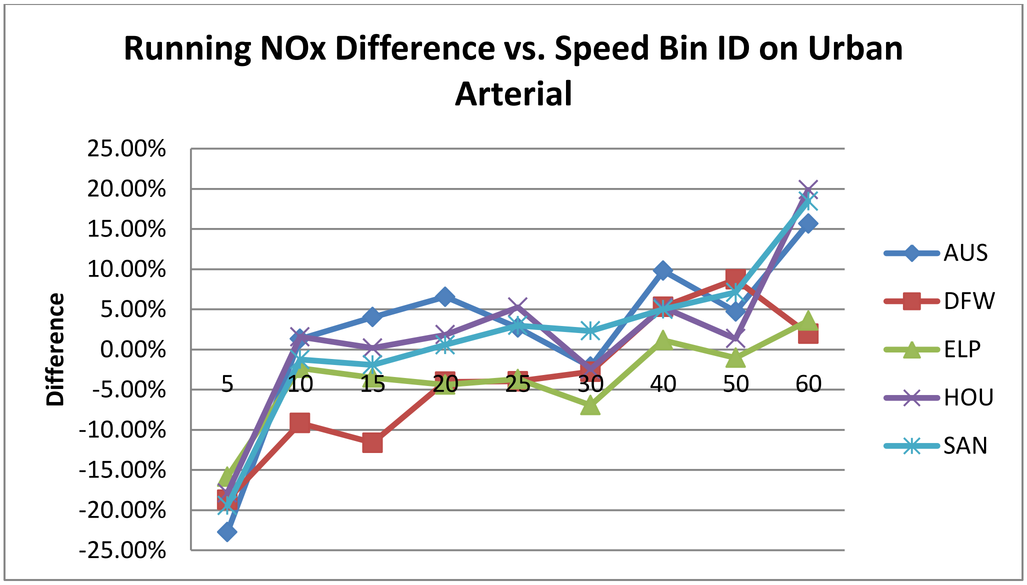
5.3.4 Running NO_X Emissions

Nitrogen oxides, mainly including nitrogen monoxide (NO) and nitrogen dioxide (NO_2), are considered as the main contribute to the acid rain or harmful photochemical smog. Figure 27 shows the percentage differences of NO_X emissions as the speed increases of all the five metropolitan areas on either urban arterials or urban freeways. Similar to the THC emission, there are several interesting facts that for urban arterials:

- The plots show an increasing trend of the differences in all five areas.
- In the case of 5 mph, the default results of MOVES overestimate the running emission rates by 17% to 23%.
- In the case of 60 mph, the default results of MOVES underestimate the emission rates up to 20%.

Also for urban freeways, the graph shows:

- The scale of the differences is smaller than the urban arterial but there is no clear similar trend shared by the five study areas.
- For DFW are and ELP area, the default results both overestimate the NO_X emissions, especially in the case of 50 mph in ELP.
- For AUS area, SAN area and HOU area, the default results all underestimate the NO_X emissions.
- There is an unusual drop when the speed goes to 70 mph from the case of 60 mph for all five study areas.



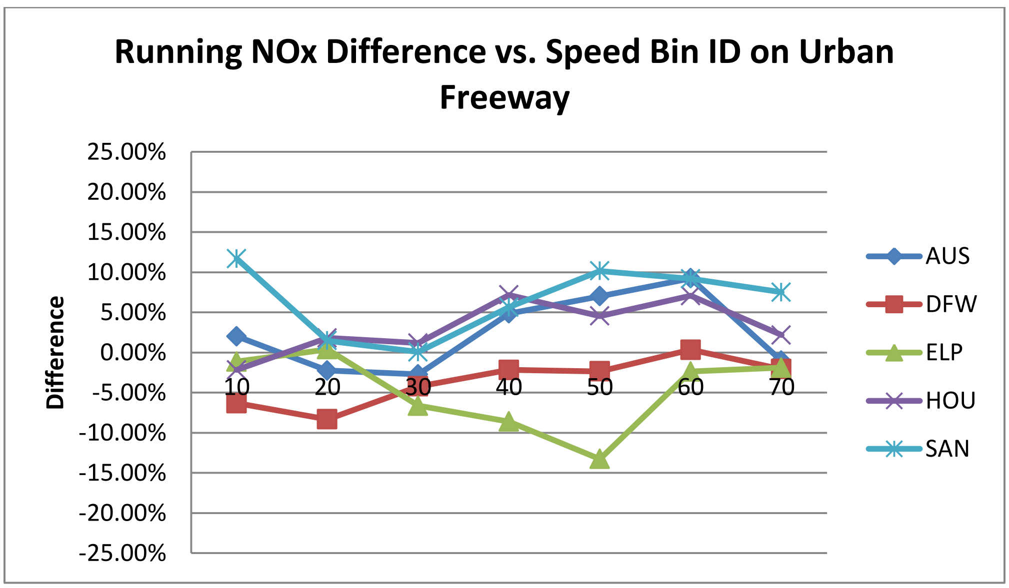


Figure 27. Distributions of NO_X Differences of Five Study Areas

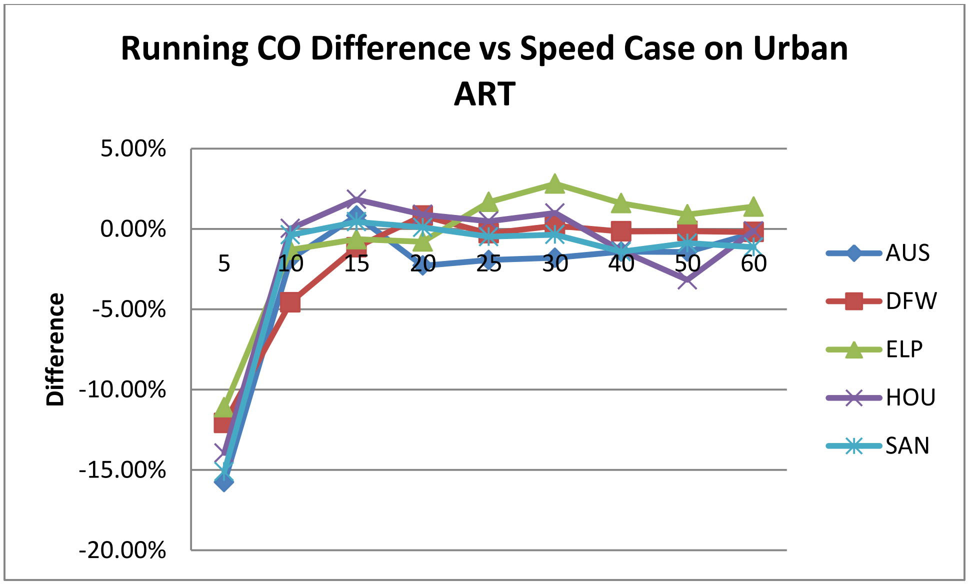
5.3.5 Running CO Emission

Different from previous two emissions, carbon monoxide indicates a part of incomplete combustion of the fuel. Figure 28 shows the percentage differences of NO_X

emissions as the speed increases of all the five metropolitan areas on either urban arterials or urban freeways. Unlike THC emission and NO_X emission, the differences are pretty small except for the case of 5 mph on urban arterials in which the default results overestimate up to 15% in the running emission rates.

5.4 Connecting Emission Rates with Drive Cycles

In previous sections, the emission rates have been successfully estimated by the inputs required by MOVES. Also, an initial analysis has been made by looking at the comparison graphs directly. As MOVES is known, the factors that would affect the estimation of emissions are drive cycles and the calculation progress. Hence, the next step of the analysis is tracing back to the drive cycles, in the format of operating mode distributions, which were developed in Section 4 since the base of the emission estimation is the database of drive cycles. As the results show in Section 5.3, some critical differences are analyzed in the subsections from drive cycle point of view. Also, as mentioned in section 4, there are two different operating mode distributions, target vector and cycle vector, and the representativeness of the latter one is the other part of the following analysis.



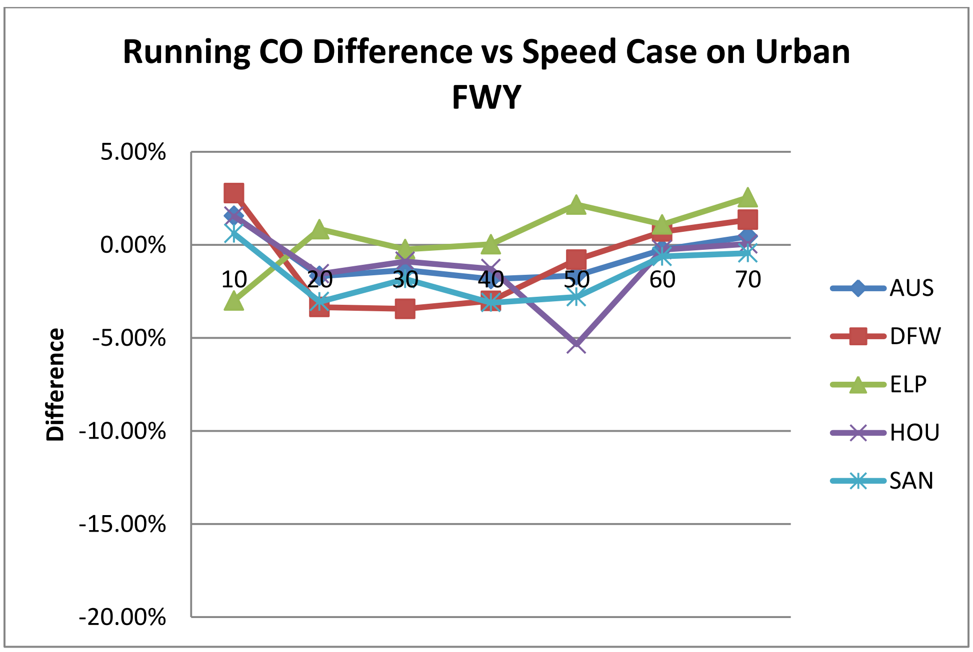


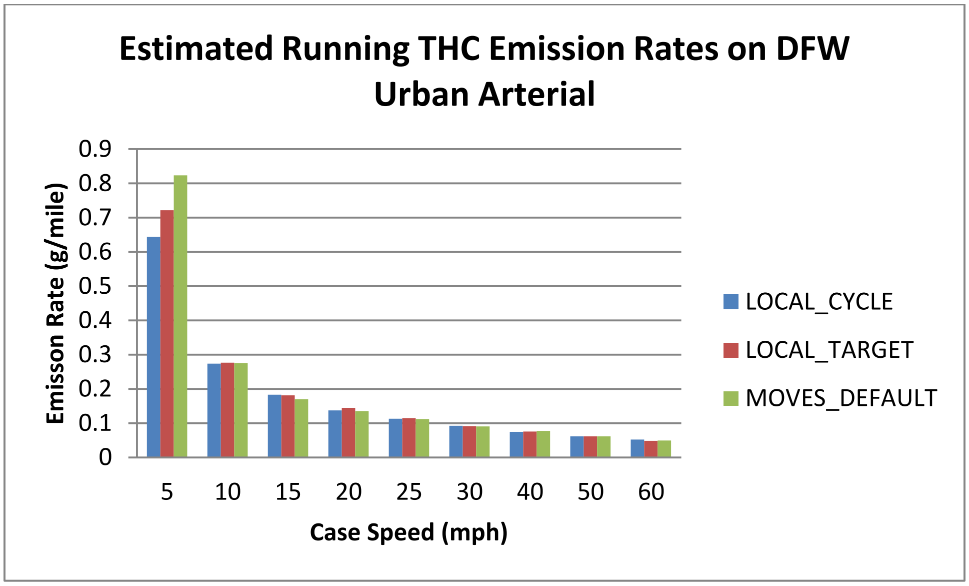
Figure 28. Distributions of CO Differences of Five Study Areas

5.4.1 The Accuracy of the Drive Cycles

As previously said, local target vector developed from local dataset is the most representative solution to present the driving patterns. However, such optimal solution could not be made to a simple speed-time trajectory graph. The local drive cycle is playing such an important role to be the sub-optimal solution while maintaining the representativeness. In previous section 4, the accuracy check is made mathematically by comparing the probability. Here, the accuracy is also checked intuitively by the graphs of the emission rates. DFW area is treated as the example to test the accuracy. Figures 29 to Figure 31 show the estimated running emission rates on all types of road types in DFW area. It is clear that the estimated emission rates from drive cycles fit those from target vector among most speed bins by looking at the chart, which is to say the cycle is enough represent the target vector to present the driving patterns and emission patterns. The divergence happens in the low speed cases, such as 5 mph on ART and 10 mph on FWY.

5.4.2 Comparative Analysis Based on Drive Cycles

Based on the graphs shown in Section 5.3, it reveals that the biggest differences lie in the case of average speed at 5 mph where the default results overestimate a lot in all the five study areas. In the opposite, there are also some cases in which the default estimates and local estimates are very close. The reason why these differences vary could be attributed to two parts, drive cycles and the calculations. For example, for the same case, the drive cycle, or the operating mode distribution, is constant while it varies among the types of emissions due to the calculation. Meanwhile, inside the estimation of one type of emission, drive cycles contribute mostly to the mutative differences among all the cases. Followed by the conclusions now, the low speed cases, regarding case of 5 mph on urban arterial as representative, and the high speed cases, regarding case of 60 mph on urban arterial as representative, are both studied from drive cycle point of view.



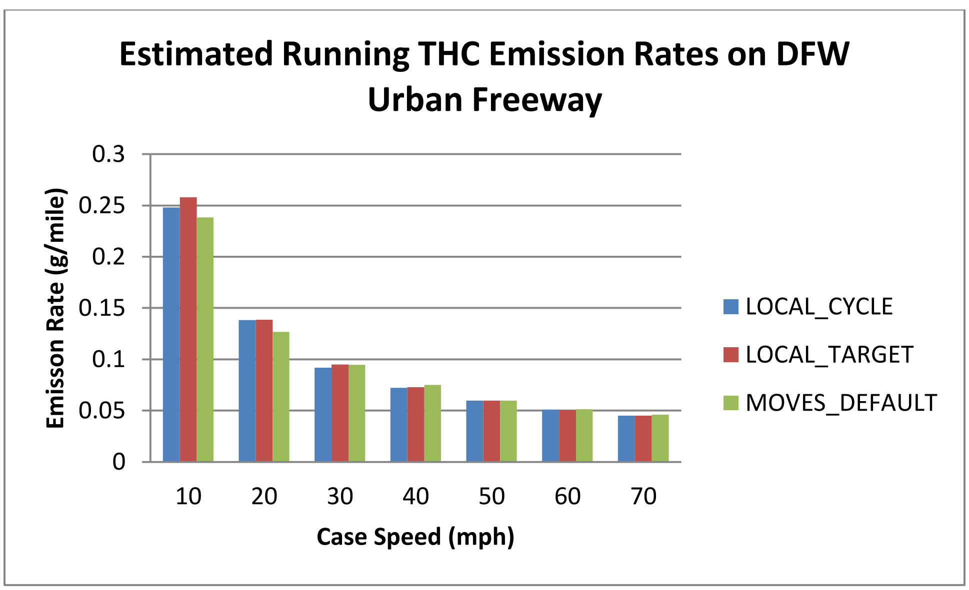
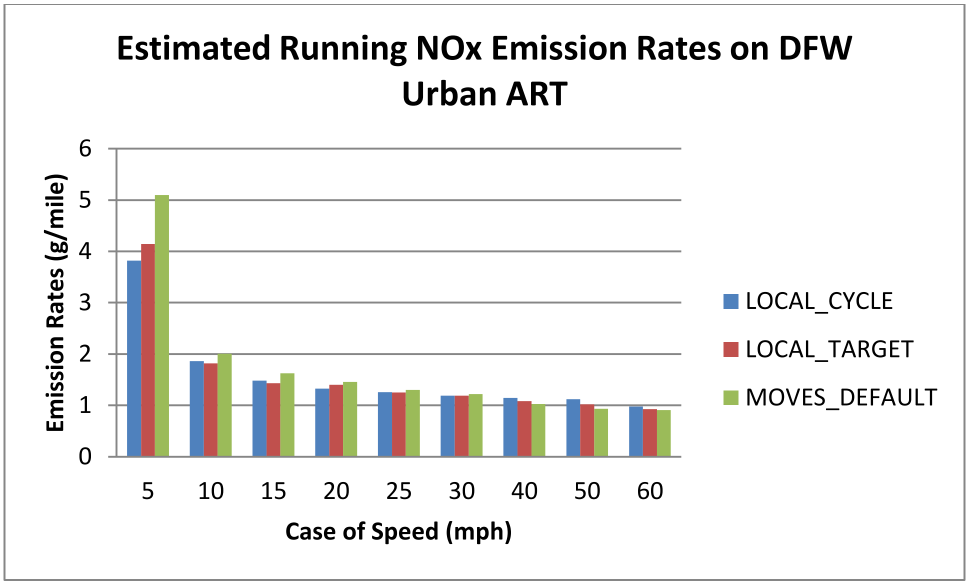


Figure 29. Estimated Running THC Emission Rates of DWF Area



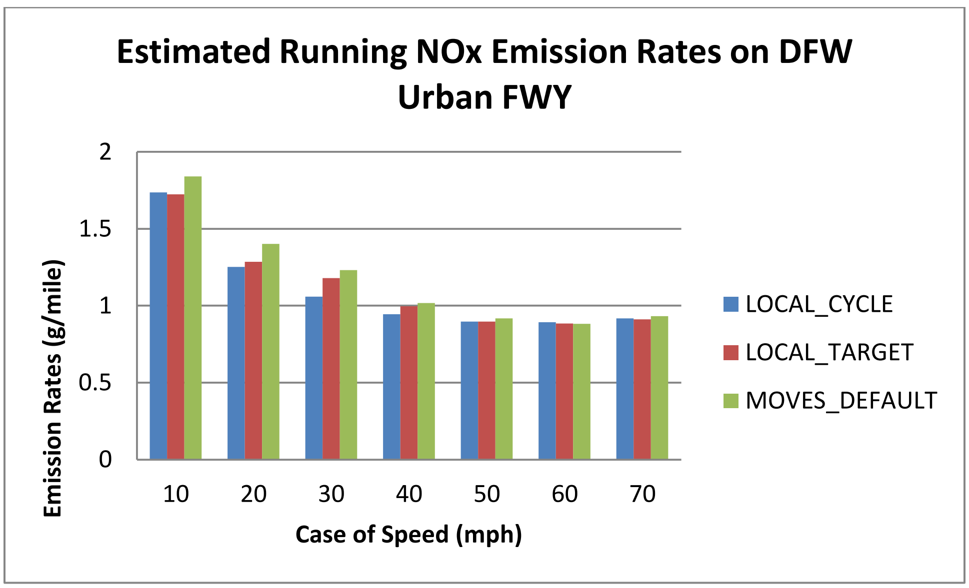
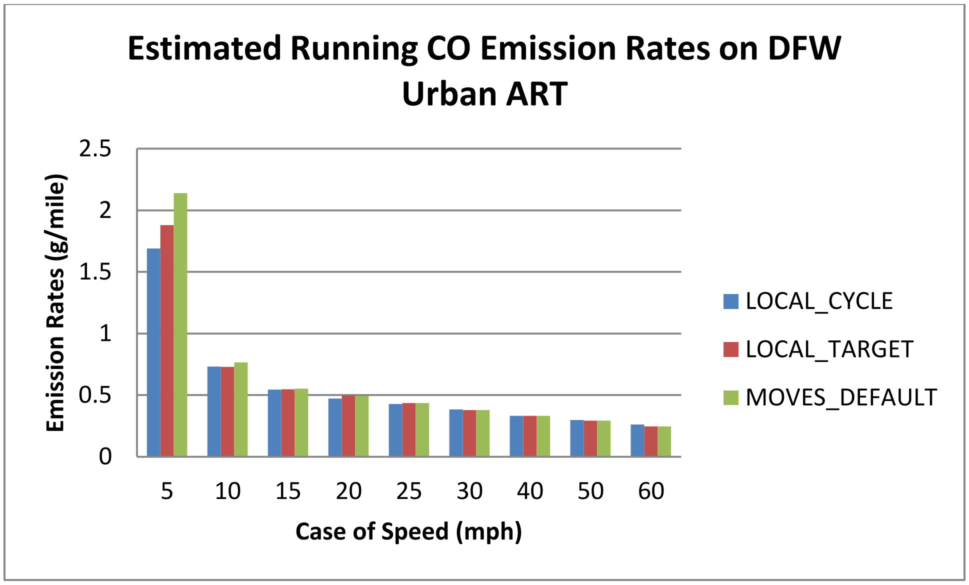


Figure 30. Estimated Running NO_X Emission Rates of DWF Area



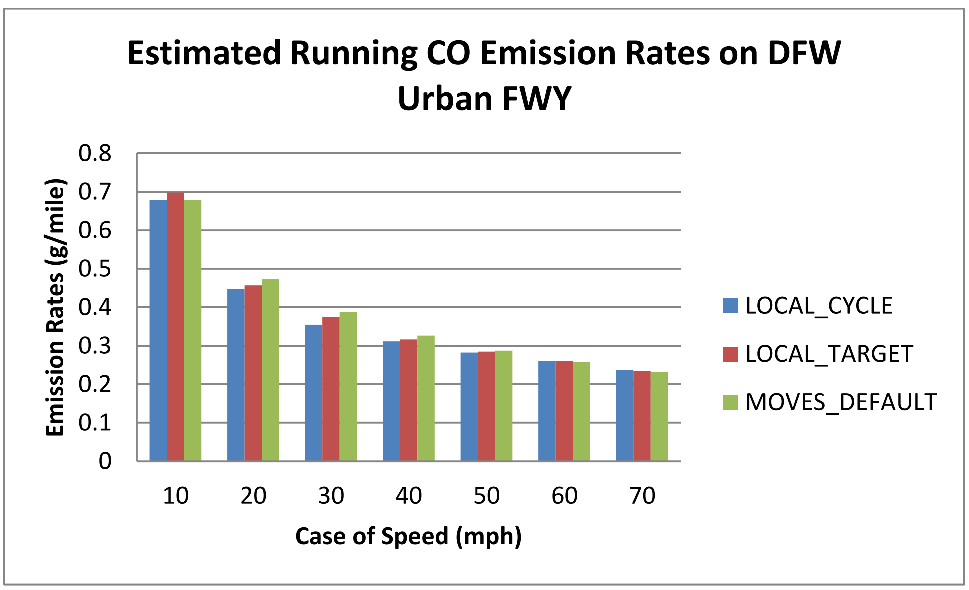


Figure 31. Estimated Running CO Emission Rates of DWF Area

Low Speed Cases (Case of 5 mph on Urban Arterials)

Regarding Austin area as an example, Figure 32 and Figure 33 provide an intuitive comparison of drive cycle as well as the operating mode distribution between local and default. It is clear that the default drive cycle has more high speed and high acceleration part than the local one which is to say the local drive cycle is somehow more constant than the default one. It is also reflected on comparison between these two operating mode distributions. The percentage of idling is as high as 41% in the local drive cycle meanwhile there is only 15% part in the default one idling. Instead, there are more shares of operating modes from 11 to 16 in default which means the more VSP and higher speed range. As it is known, MOVES runs a running emission process to the drive cycles in which the idling part generates less emissions than the acceleration process, especially when speeds are low there are a lot of start situations. And the fact fits the emission estimation graphs shown above. At this point, MOVES' default drive cycles underestimate the idling part of such low speed cases which leads to the large differences shown.

High Speed Cases (Case of 60 mph on Urban Arterials)

Unlike the low speed cases, the high speed cases on the urban arterial are not exist in the MOVES database. There is an assumption that MOVES apply the cases with the average speed closest to the input of average speed. Therefore the case that MOVES selected for case at 60 mph is the case that developed on freeway. Logically, the speed variation is larger on the arterial than on the freeway. Figure 34 and Figure 35 put some evidences on the suspect. From the operating mode distributions, MOVES' drive cycle concentrate most observations in operating mode 35 while the operating mode distribution is more spread out which means there are higher speeds and more frequently speed changes. From the figure 34, it has been proved the conclusion from the inference from the operating mode distribution. With such big differences, on both drive cycles and the emission rates, especially NO_X emissions, there is a requirement to create the default cases into MOVES default database for the data coming from high speed cases on the arterial.

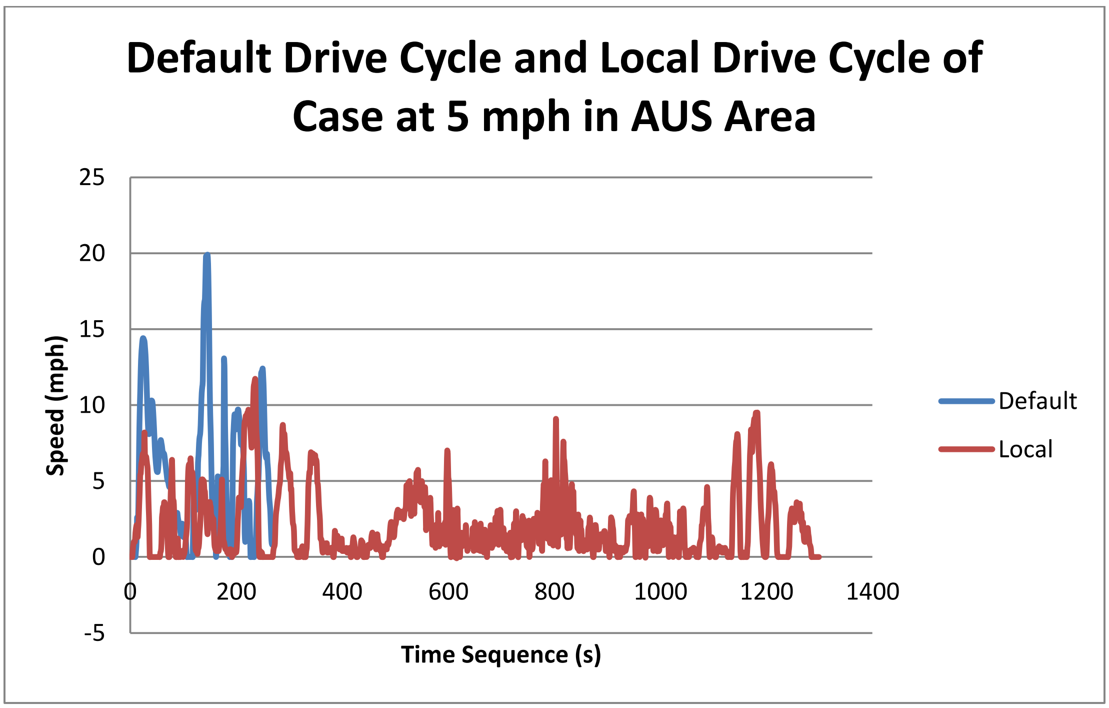


Figure 32. Drive Cycle Comparison Plot of Case at 5 mph in AUS Area

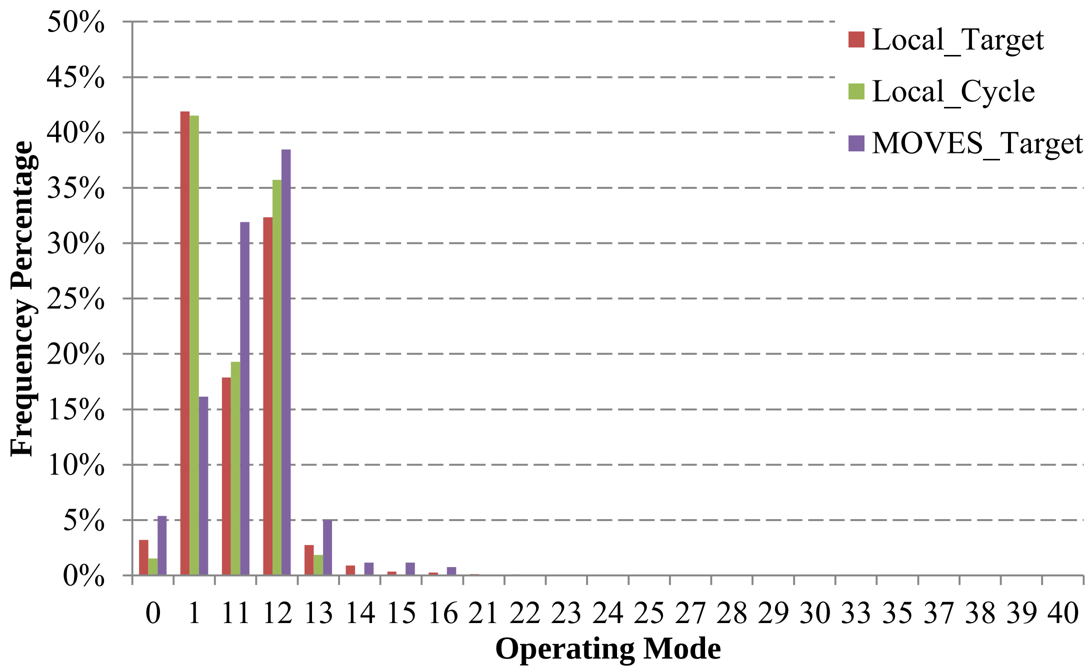


Figure 33. Operating Mode Distributions of Case at 5 mph in AUS Area

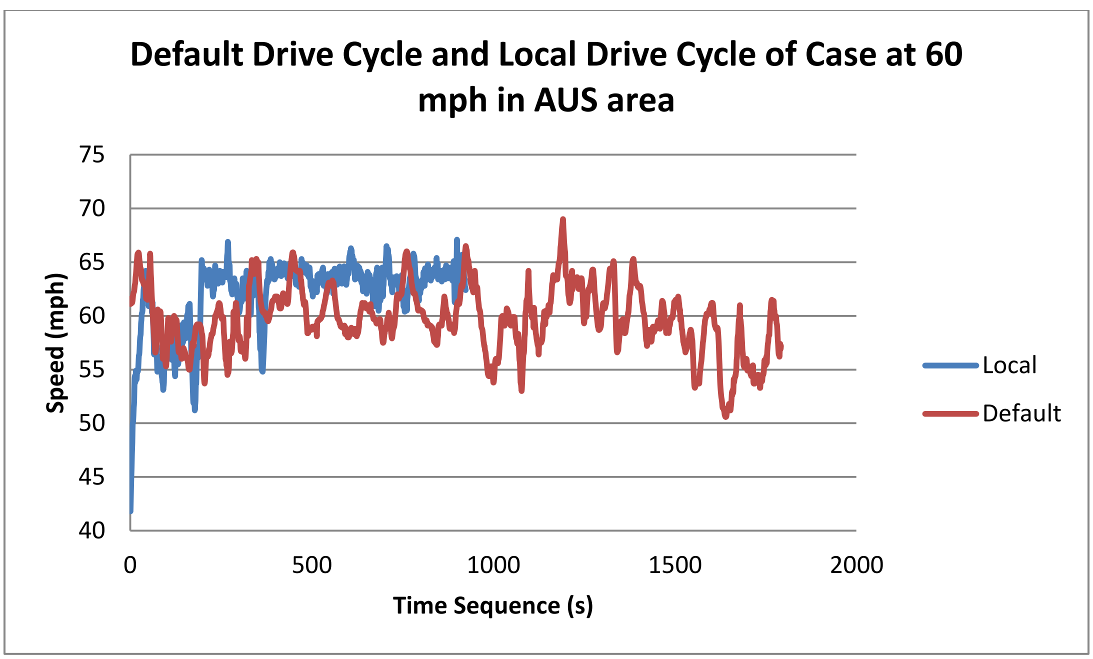


Figure 34. Drive Cycle Comparison Plot of Case at 60 mph in AUS Area

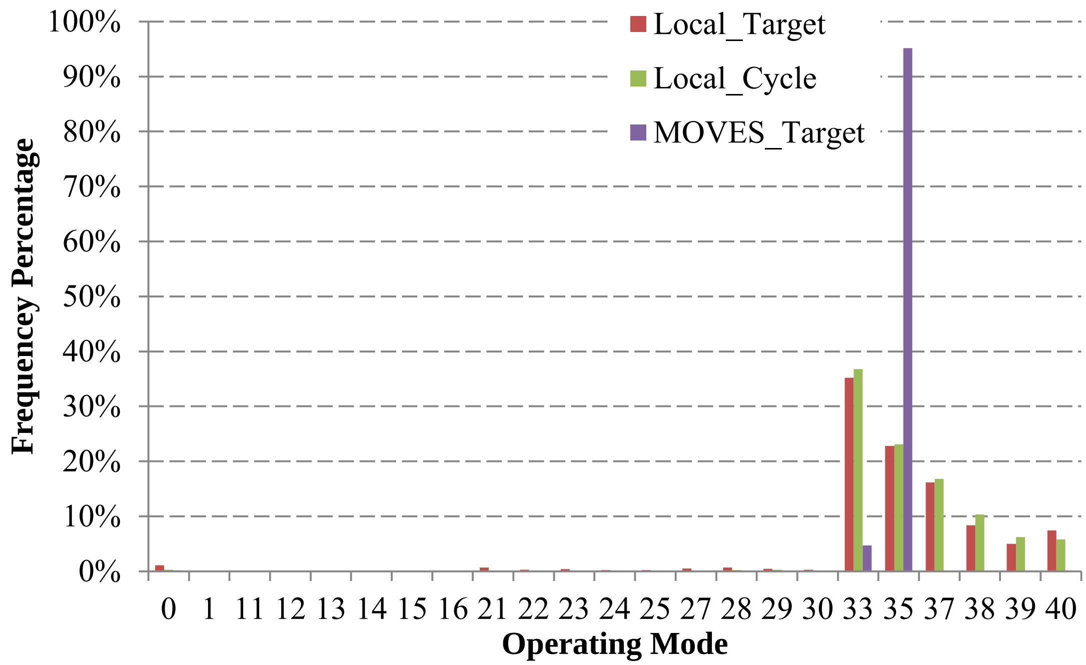


Figure 35. Operating Mode Distributions of Case at 60 mph in AUS Area

6. RECOMMENDATION AND FUTURE WORK

6.1 Summary

In this study, several objectives have been accomplished. First of all, the data collection method, GPS data logging, has been tested on accuracy by comparing with the OBD method, which validates the foundation of the research. Also, based on the methodology proposed by EPA, the progress of the raw data is modified to adjust to local characteristics. New qualifying thresholds are added when tearing apart the whole dataset to micro-trips which is the base of building up the drive cycles. Followed by the initial processing, local target vector and drive cycles are gradually developed in the format of operating mode distributions. All of these pre-processing paves a good way for the emission estimation and comparison. With the help by MOVES, the moving information is translated to the estimation of the rates of selected emissions. Two different comparisons are made afterwards. One is the difference between the emission rates from local target vectors and the default emission rates MOVES comes with. The other one is the representativeness of local drive cycles compared to the local target vector.

6.2 Findings and Limitations

The empirical studies in the thesis research have shown the importance of local drive cycles in emission estimation by MOVES. All findings are listed below:

- It is important to take care of the initial processing of the raw data. The methodology that applied in data cleansing and initial processing has significant impact on building up the micro-trips. Different thresholds lead to different driving patterns inside of the micro-trips so that the emission estimation will be more realistic.

- While it also draws the importance of developing local drive cycles by looking at the comparisons of the emission rates especially for the low speed cases of study area. MOVES underestimate the idling part in the extreme low cases in which the drive cycles and the estimated emission rates have significant differences.
- The emission estimation of the MOVES is still accurate for mid-speed cases, at magnitude level. There are not big differences between local drive cycles and default drive cycles. Hence, the results generated from the research is kind of validation of MOVES' calculation.
- Also, the study provides an initial work on completing the drive cycle databases in the MOVES. The results suggest the requirement of developing drive cycles for HDDVs at the high speed over 30 mph on arterials. For the low speed cases on the freeways, the differences are significant but not that critical.

It is said that nothing in the world is perfect. Therefore, other than the findings above, there are some limitations for the emission rates as follows:

- Further comparisons are still required to made between on-road emissions and estimated emissions to analyze the accuracy of the estimated emissions to the emissions when vehicle is moving in real world.

- The drive cycles are developed in the format of operating mode distributions. While the original drive cycle is in second-by-second format. There might be some differences when applying either of these two in emission estimation which has not been tested in this research.

6.3 Future Work

As scheduled, the outcomes generated by MOVES will be compared to in-use results. A PEMS (SEMTECH-DS) will be used for second-by-second collection of gaseous pollutants (including NOx and CO) and a DMM will be used to measure PM emissions. For DMM measurements, diluted samples will be drawn through micro dilution sampling system. Filter samples will be drawn through the micro dilution sampling system, and the collected filters and cartridges will be sent to Oak Ridge National Laboratory for PM mass and MSAT analysis. MSATs that will be analyzed are acetaldehyde, formaldehyde, acrolein. This will provide the opportunity to compare in-use emissions measured by PEMS equipment with those produced by MOVES for the same driving and ambient conditions. The comparison between the results of observed in-use emissions and MOVES estimates can be performed in two ways: aggregate level, and disaggregate (VSP/speed point comparison). In the aggregate level comparison, the observed driving schedules (second-by-second speed profile) of a test vehicle is input to MOVES. The total emissions associated with the period of the drive schedule are then estimated by the model. These results can be then transformed into average distance-based emission rates; i.e. g/mi. This estimate then will be compared to the total measured emissions or corresponding distance-based emissions obtained from field observations.

REFERENCES

1. **Kite, Chris.** Preliminary Comparison Between MOVES and MOBILE 6. [Online] http://www.tceq.texas.gov/assets/public/implementation/air/am/committees/pmt_set/20110224/20110224-kite.pdf. Februray, 2011.

2. **Carleson.R Thomas, Thomas Austin, Dennis McClement, and Sun-Hoon Yoon.** *Development of Gneric Link-Level Driving Cycles.* Washington, D.C. : Prepared for U.S. Environmental Protection Agency, 2009.

3. **Yu Lei, Ziqianli Wang, and Qinyi Shi.** *PEMS-Based Approach to Developing and Evaluating Driving Cycles for Air Quality Assessment.* The Texas A&M University System : Texas Transportation Institute, 2010. Report 169300-1.

4. **EPA.** 1997 Development of Speed Correction Cycles. [Online] http://www.epa.gov/otaq/models/mobile6/m6spd001.pdf. SR97-04-01.

5. **Moorey J., T.Limanond, and D. Niermeier.** *Validity of Chase Car Data Used in Developing Emissions Cycles.* Journal of Transportation and Statistics, September 2000, Vol. 3. 2.

6. **Battelle Transportation Division.** *Heavy-Duty Truck Activity Data.* Federal Highway Administration: Office of Highway Information Management, Washington, D.C., 1999

7. **Montazeri-Gh, M. and M. Naghizadeh.** *Development of Car Drive Cycle for Simulation of Emissions and Fuel Economy.* Proceeding of the 15th European

Simulation Symposium, Alexander Verbraeck, Vlatka Hlupic (Eds.) SCS European Council/SCS Europe BVBA, 535-540.

8. **Jackson, E., L. Aultman-Hall, B. Holmen, and J. Du.** *Evaluating the Ability of Global Positioning System Receivers to Measure a Real-World Operating Mode for Emissions Research.* Transportation Research Record: Transportation Research Board, Washington, D.C., 2005.

9. **Alessandrini, A. and F. Orecchini.** *A Driving Cycle for Electrically-Driven Vehicles in Rome.* 9, Institution of Mechanical Engineers, Part D: Journal of Automobile Enineering, 2003, Vol. 217. 781-789.

10. **EPA.** 2009 MOVES 2009 Software Deisgn and Reference Manual. [Online] http://www.epa.gov/otaq/models/moves/420b09007.pdf.

11. **Andre, M., R. Jourmard, R. Vidon, P.Tassel and P.Perret.** *Real-World European Driving Cycles, for Measuring Pollutant Emissions from High- and Low-Powered Cars*, Atmospheric Environment, 2006, Vol. 40, 5944-5953.

12. **Ergeneman C. Sorusbay, and A. Goktan M.** *Development of a Driving Cycle for the Prediction of Pollutant Emissions and Fuel Comsumption.* 391-399, International Journal of Vehicle Design, 1997, Vol. 18.

13. **EPA.** 2009 MOVES 2010 User Guide. [Online] http://www.epa.gov/otaq/models/moves/420b09041.pdf. September, 2009.

14. **Niemeier, D.A., Thirayoot Limanond, and Jennifer E. Morey.** *Data Collection for Driving Cycle Development: Evaluation of Data Collection Protocols,* University of California, Davis: Institute of Transportation Studies, 1999.

15. **Federal Highway Administration.** HPMS Field Manual: Chapter 3: Summary Data Requirements. [Online] http:\\www.fhwa.dot.gov/ohim/hpmsmanl/chapt3.cfm. September, 2010.

APPENDIX A: CONSENT FORM

Introduction

The purpose of this form is to provide you information that may affect your decision as to whether or not to participate in this research study. If you decide to participate in this study, this form will also be used to record your consent.

You have been asked to participate in a research project studying travel behaviors in your region and how emissions predicted from this behavior compare to actual emissions measured in a region. The purpose of this study is to gather the necessary real-world speed data for various vehicle types to run the modeling software. You were selected to be a possible participant because you responded to a posted recruitment notice. This study is being sponsored or funded by the Texas Department of Transportation (TxDOT).

What will I be asked to do?

If you agree to participate in this study, you will be asked to carry a set of GPS units in your vehicle for seven-to-thirty (7-to-30) days to identify the type of roads you travel on, collecting speed and location data while you travel as you normally would during that time.

What are the risks involved in this study?

The risks associated with this study are that researchers will be able to determine the destinations of your daily travel activity during data collection. Other risks associated in this study are minimal, and are not greater than risks ordinarily encountered in daily life.

What are the possible benefits of this study?

You will receive no direct benefit from participating in this study; however, your participation will provide TxDOT the data needed to assess the accuracy of air quality predictions provided by the Environmental Protection Agency's MOVES emissions model.

Do I have to participate?

No. Your participation is voluntary. You may decide not to participate or to withdraw at any time without your current or future relations with the Texas Transportation Institute, The Texas A&M University System, or the Texas Department of Transportation being affected.

Will I be compensated?

The following groups are not eligible for compensation:

- Employees and their family members of the research organization: Texas Transportation Institute (TTI); and

- Employees and their family members of beneficiary agencies including TxDOT, TCEQ, NCTCOG, ACOG, City of Houston, City of Austin, HGAC.

Eligible participants will receive seventy five dollars ($75) in cash for participating in the study. Disbursement will occur when you return the GPS unit to the researchers and after researchers validate the data at the end of the study period. In the event of data failure or errors, you might be asked to continue your participation for another round of data collection with another set of data loggers.

Who will know about my participation in this research study?

This study is confidential, and all location data gathered will be stored as encrypted data files. Data files will be labeled in such a way that individual file names cannot be linked back to specific participants. The records of this study will be kept private. No identifiers linking you to this study will be included in any sort of report that might be published. Research records will be stored securely and only the research team will have access to the records.

Whom do I contact with questions about the research?

If you have questions regarding this study, you may contact Mohamadreza Farzaneh, 512-467-0946, mfarzaneh@tamu.edu.

Whom do I contact about my rights as a research participant?

This research study has been reviewed by the Human Subjects' Protection Program and/or the Institutional Review Board at Texas A&M University. For research-related problems or questions regarding your rights as a research participant, you can contact these offices at (979)458-4067 or irb@tamu.edu.

Signature

Please be sure you have read the above information, asked questions and received answers to your satisfaction. You will be given a copy of the consent form for your records. By signing this document, you consent to participate in this study.

Signature of Participant: ____________________________Date: ______________

Printed Name: ____

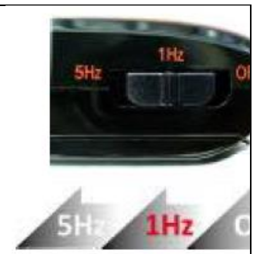
Signature of Person Obtaining Consent: _______________Date: ______________

Printed Name:____

APPENDIX B: INSTALLATION INSTRUCTIONS FOR GPS UNITS

1) You must have received three GPS units that are tied together. If you have not received three (3) units, immediately contact TTI researchers at (512) 467-0946 or mfarzaneh@tamu.edu.

2) Turn on the GPS units by pushing the mode switch located on the side of the unit to the right. Put the mode switch on **<u>1Hz</u>** (middle position). The 5Hz option (leftmost position) is blocked on all units.

<table>
<tr>
<td></td>
<td>Mode Switch on
1Hz</td>
<td>
</td>
</tr>
</table>

3) Confirm that ***ALL*** the units are functioning – both a blue symbol and an orange symbol should light up. The orange symbol will blink when receiving information from the GPS satellite system. The units will beep when they first acquire GPS location data from the system.
4) Place the active GPS unit in the driver-side door storage compartment of your vehicle, taking care to place it in such a way that it will not become buried under other items in the compartment.

If your vehicle does not have a door storage compartment, place the units in the glove box or a secure place inside the cabin of the vehicle. The units should be placed in a way that does not pose any danger to the driver; e.g., block the driver's view, etc.

5) Leave units in the “ON – 1Hz” position and in the secure location until the end of the research experiment period.
6) Upon completion of the research experiment, remove the GPS units from your vehicle, turn off the unit, and return the unit to the researcher in the pre-addressed box that was provided with the unit.

Return Address:

Texas Transportation Institute

Center for Air Quality Studies

1106 Clayton Lane, Suite 300E

Austin, TX 78723

(512) 467-0946